PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/82, 9/00, A01H 5/00, C12Q 1/68 | A2

(11) International Publication Number: **WO 00/15809**

(43) International Publication Date: 23 March 2000 (23.03.00)

(21) International Application Number: PCT/EP99/06766

(22) International Filing Date: 13 September 1999 (13.09.99)

(30) Priority Data:
09/153,278 15 September 1998 (15.09.98) US

(71) Applicant *(for all designated States except AT US)*: NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH–4058 Basel (CH).

(71) Applicant *(for AT only)*: NOVARTIS–ERFINDUNGEN VERWALTUNGSGESELLSCHAFT M.B.H. [AT/AT]; Brunner Strasse 59, A–1230 Vienna (AT).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LEVIN, Joshua, Zvi [US/US]; 1008 Urban Avenue, Durham, NC 27701 (US). POTTER, Sharon, Lee [US/US]; 3837 Whispering Branch Road, Raleigh, NC 27613 (US). BUDZISZEWSKI, Gregory, Joseph [US/US]; 2016 Englewood Avenue, Durham, NC 27705 (US). WARD, Eric, Russell [US/US]; 3761 Bentley Drive, Durham, NC 27707 (US). BERNASCONI, Paul [US/US]; 102 Windhover Drive, Chapel Hill, NC 27514 (US).

(74) Agent: BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent & Trademark Dept., CH–4002 Basel (CH).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: URACIL PERMEASE FROM ARABIDOPSIS AS HERBICIDAL TARGET GENE

(57) Abstract

The invention relates to a gene isolated from *Arabidopsis* that codes for a protein essential for seedling growth. The invention also includes the methods of using this protein to discover new herbicides, based on the essentiality of the gene for normal growth and development. The invention can also be used in a screening assay to identify inhibitors that are potential herbicides. The invention is also applied to the development of herbicide tolerant plants, plant tissues, plant seeds, and plant cells.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

URACIL PERMEASE FROM ARABIDOPSIS AS HERBICIDAL TARGET GENE

The invention relates to a gene isolated from *Arabidopsis* that codes for a protein essential for seedling growth. The invention also includes the methods of using this protein as an herbicide target, based on the essentiality of the gene for normal growth and development. The invention is also useful as a screening assay to identify inhibitors that are potential herbicides. The invention may also be applied to the development of herbicide tolerant plants, plant tissues, plant seeds, and plant cells.

The use of herbicides to control undesirable vegetation such as weeds in crop fields has become almost a universal practice. The herbicide market exceeds 15 billion dollars annually. Despite this extensive use, weed control remains a significant and costly problem for farmers.

Effective use of herbicides requires sound management. For instance, the time and method of application and stage of weed plant development are critical to getting good weed control with herbicides. Since various weed species are resistant to herbicides, the production of effective new herbicides becomes increasingly important. Novel herbicides can now be discovered using high-throughput screens that implement recombinant DNA technology. Metabolic enzymes found to be essential to plant growth and development can be recombinantly produced through standard molecular biological techniques and utilized as herbicide targets in screens for novel inhibitors of the enzyme activity. The novel inhibitors discovered through such screens may then be used as herbicides to control undesirable vegetation.

Herbicides that exhibit greater potency, broader weed spectrum, and more rapid degradation in soil can also, unfortunately, have greater crop phytotoxicity. One solution applied to this problem has been to develop crops that are resistant or tolerant to herbicides. Crop hybrids or varieties tolerant to the herbicides allow for the use of the herbicides to kill weeds without attendant risk of damage to the crop. Development of tolerance can allow application of a herbicide to a crop where its use was previously precluded or limited (*e.g.* to pre-emergence use) due to sensitivity of the crop to the herbicide. For example, U.S. Patent No. 4,761,373 to Anderson *et al.* is directed to plants resistant to various imidazolinone or sulfonamide herbicides. The resistance is conferred by an altered acetohydroxyacid synthase (AHAS) enzyme. U.S. Patent No. 4,975,374 to Goodman *et al.* relates to plant cells and plants containing a gene encoding a mutant

glutamine synthetase (GS) resistant to inhibition by herbicides that were known to inhibit GS, *e.g.* phosphinothricin and methionine sulfoximine. U.S. Patent No. 5,013,659 to Bedbrook *et al.* is directed to plants expressing a mutant acetolactate synthase that renders the plants resistant to inhibition by sulfonylurea herbicides. U.S. Patent No. 5,162,602 to Somers *et al.* discloses plants tolerant to inhibition by cyclohexanedione and aryloxyphenoxypropanoic acid herbicides. The tolerance is conferred by an altered acetyl coenzyme A carboxylase (ACCase).

Notwithstanding the above described advancements, there remains persistent and ongoing problems with unwanted or detrimental vegetation growth (e.g. weeds). Furthermore, as the population continues to grow, there will be increasing food shortages. Therefore, there exists a long felt, yet unfulfilled need, to find new, effective, and economic herbicides.

DEFINITIONS

For clarity, certain terms used in the specification are defined and presented as follows:

Chimeric: is used to indicate that a DNA sequence, such as a vector or a gene, is comprised of more than one DNA sequences of distinct origin which are fused together by recombinant DNA techniques resulting in a DNA sequence, which does not occur naturally,

Co-factor: natural reactant, such as an organic molecule or a metal ion, required in an enzyme-catalyzed reaction. A co-factor is e.g. NAD(P), riboflavin (including FAD and FMN), folate, molybdopterin, thiamin, biotin, lipoic acid, pantothenic acid and coenzyme A, S-adenosylmethionine, pyridoxal phosphate, ubiquinone, menaquinone. Optionally, a co-factor can be regenerated and reused.

DNA shuffling: DNA shuffling is a method to introduce mutations or rearrangements, preferably randomly, in a DNA molecule or to generate exchanges of DNA sequences between two or more DNA molecules, preferably randomly. The DNA molecule resulting from DNA shuffling is a shuffled DNA molecule that is a non-naturally occurring DNA molecule derived from at least one template DNA molecule. The shuffled DNA encodes an enzyme modified with respect to the enzyme encoded by the template DNA, and preferably has an altered biological activity with respect to the enzyme encoded by the template DNA.

Enzyme activity: means herein the ability of an enzyme to catalyze the conversion of a substrate into a product. A substrate for the enzyme comprises the natural substrate of the enzyme but also comprises analogues of the natural substrate which can also be

converted by the enzyme into a product or into an analogue of a product. The activity of the enzyme is measured for example by determining the amount of product in the reaction after a certain period of time, or by determining the amount of substrate remaining in the reaction mixture after a certain period of time. The activity of the enzyme is also measured by determining the amount of an unused co-factor of the reaction remaining in the reaction mixture after a certain period of time or by determining the amount of used co-factor in the reaction mixture after a certain period of time. The activity of the enzyme is also measured by determining the amount of a donor of free energy or energy-rich molecule (e.g. ATP, phosphoenolpyruvate, acetyl phosphate or phosphocreatine) remaining in the reaction mixture after a certain period of time or by determining the amount of a used donor of free energy or energy-rich molecule (e.g. ADP, pyruvate, acetate or creatine) in the reaction mixture after a certain period of time.

Expression: refers to the transcription and/or translation of an endogenous gene or a transgene in plants. In the case of antisense constructs, for example, expression may refer to the transcription of the antisense DNA only.

Gene: refers to a coding sequence and associated regulatory sequences wherein the coding sequence is transcribed into RNA such as mRNA, rRNA, tRNA, snRNA, sense RNA or antisense RNA. Examples of regulatory sequences are promoter sequences, 5' and 3' untranslated sequences. Further elements that may be present are, for example, introns.

...gene of interest: refers to any gene which, when transferred to a plant, confers upon the plant a desired characteristic such as antibiotic resistance, virus resistance, insect resistance, disease resistance, or resistance to other pests, herbicide tolerance, improved nutritional value, improved performance in an industrial process or altered reproductive capability. The "gene of interest" may also be one that is transferred to plants for the production of commercially valuable enzymes or metabolites in the plant.

Herbicide: a chemical substance used to kill or suppress the growth of plants, plant cells, plant seeds, or plant tissues.

Heterologous DNA Sequence: a DNA sequence not naturally associated with a host cell into which it is introduced, including non-naturally occurring multiple copies of a naturally occurring DNA sequence.

Homologous DNA Sequence: a DNA sequence naturally associated with a host cell into which it is introduced.

Identity: The percentage of sequence identity is determined using computer programs that are based on dynamic programming algorithms. Computer programs that are

preferred within the scope of the present invention include the BLAST (Basic Local Alignment Search Tool) search programs designed to explore all of the available sequence databases regardless of whether the query is protein or DNA. Version BLAST 2.0 (Gapped BLAST) of this search tool has been made publicly available on the Internet (currently http://www.ncbi.nlm.nih.gov/BLAST/). It uses a heuristic algorithm which seeks local as opposed to global alignments and is therefore able to detect relationships among sequences which share only isolated regions. The scores assigned in a BLAST search have a well-defined statistical.

Inhibitor: a chemical substance that inactivates the enzymatic activity of a protein such as a biosynthetic enzyme, receptor, signal transduction protein, structural gene product, or transport protein that is essential to the growth or survival of the plant. In the context of the instant invention, an inhibitor is a chemical substance that inactivates the enzymatic activity of 4788 from a plant.

Isogenic: plants which are genetically identical, except that they may differ by the presence or absence of a transgene.

Isolated: in the context of the present invention, an isolated DNA molecule or an isolated enzyme is a DNA molecule or enzyme that, by the hand of man, exists apart from its native environment and is therefore not a product of nature. An isolated DNA molecule or enzyme may exist in a purified form or may exist in a non-native environment such as, for example, a transgenic host cell.

Marker gene: a gene encoding a selectable or screenable trait.

Mature protein: protein which is normally targeted to a cellular organelle, such as a chloroplast, and from which the transit peptide has been removed.

Minimal Promoter: promoter elements, particularly a TATA element, that are inactive or that have greatly reduced promoter activity in the absence of upstream activation. In the presence of a suitable transcription factor, the minimal promoter functions to permit transcription.

Modified Enzyme Activity: enzyme activity different from that which naturally occurs in a plant (i.e. enzyme activity that occurs naturally in the absence of direct or indirect manipulation of such activity by man), which is tolerant to inhibitors that inhibit the naturally occurring enzyme activity.

Operably linked to/ associated with: a regulatory DNA sequence is said to be "operably linked to" or "associated with" a DNA sequence that codes for an RNA or a

protein if the two sequences are situated such that the regulatory DNA sequence affects expression of the coding DNA sequence.

Plant: refers to any plant, particularly to seed plants

Plant cell: structural and physiological unit of the plant, comprising a protoplast and a cell wall. The plant cell may be in form of an isolated single cell or a cultured cell, or as a part of higher organized unit such as, for example, a plant tissue, or a plant organ.

Plant material: refers to leaves, stems, roots, flowers or flower parts, fruits, pollen, pollen tubes, ovules, embryo sacs, egg cells, zygotes, embryos, seeds, cuttings, cell or tissue cultures, or any other part or product of a plant.

Pre-protein: protein which is normally targeted to a cellular organelle, such as a chloroplast, and still comprising its transit peptide.

Recombinant DNA molecule: a combination of DNA sequences that are joined together using recombinant DNA technology

recombinant DNA technology: procedures used to join together DNA sequences as described, for example, in Sambrook et al., 1989, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Selectable marker: a gene whose expression in a plant cell gives the cell a selective advantage. The selective advantage possessed by the cells transformed with the selectable marker gene may be due to their ability to grow in the presence of a negative selective agent, such as an antibiotic or a herbicide, compared to the growth of non-transformed cells. The selective advantage possessed by the transformed cells, compared to non-transformed cells, may also be due to their enhanced or novel capacity to utilize an added compound as a nutrient, growth factor or energy source. Selectable marker gene also refers to a gene or a combination of genes whose expression in a plant cell gives the cell both, a negative and a positive selective advantageSignificant Increase: an increase in enzymatic activity that is larger than the margin of error inherent in the measurement technique, preferably an increase by about 2-fold or greater of the activity of the wild-type enzyme in the presence of the inhibitor, more preferably an increase by about 5-fold or greater, and most preferably an increase by about 10-fold or greater.

Significantly less: means that the amount of a product of an enzymatic reaction is larger than the margin of error inherent in the measurement technique, preferably a decrease by about 2-fold or greater of the activity of the wild-type enzyme in the absence of the inhibitor, more preferably an decrease by about 5-fold or greater, and most preferably an decrease by about 10-fold or greater.

In its broadest sense, the term "substantially similar", when used herein with respect to a nucleotide sequence, means a nucleotide sequence corresponding to a reference nucleotide sequence, wherein the corresponding sequence encodes a polypeptide having substantially the same structure and function as the polypeptide encoded by the reference nucleotide sequence, e.g. where only changes in amino acids not affecting the polypeptide function occur. Desirably the substantially similar nucleotide sequence encodes the polypeptide encoded by the reference nucleotide sequence. The term "substantially similar" is specifically intended to include nucleotide sequences wherein the sequence has been modified to optimize expression in particular cells. The percentage of identity between the substantially similar nucleotide sequence and the reference nucleotide sequence desirably is at least 65%, more desirably at least 75%, preferably at least 85%, more preferably at least 90%, still more preferably at least 95%, yet still more preferably at least 99%. Sequence comparisons are carried out using a Smith-Waterman sequence alignment algorithm (see e.g. Waterman, M.S. Introduction to Computational Biology: Maps, sequences and genomes. Chapman & Hall. London: 1995. ISBN 0-412-99391-0, or at **http://www-hto.usc.edu/software/seqaln/index.html**). The localS program, version 1.16, is used with following parameters: match: 1, mismatch penalty: 0.33, open-gap penalty: 2, extended-gap penalty: 2. A nucleotide sequence "substantially similar" to reference nucleotide sequence hybridizes to the reference nucleotide sequence in 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$, 1 mM EDTA at 50°C with washing in 2X SSC, 0.1% SDS at 50°C, more desirably in 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$, 1 mM EDTA at 50°C with washing in 1X SSC, 0.1% SDS at 50°C, more desirably still in 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$, 1 mM EDTA at 50°C with washing in 0.5X SSC, 0.1% SDS at 50°C, preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$, 1 mM EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 50°C, more preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$, 1 mM EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 65°C.

The term "substantially similar", when used herein with respect to a protein, means a protein corresponding to a reference protein, wherein the protein has substantially the same structure and function as the reference protein, e.g. where only changes in amino acids sequence not affecting the polypeptide function occur. When used for a protein or an amino acid sequence the percentage of identity between the substantially similar and the reference protein or amino acid sequence desirably is at least 65%, more desirably at least

75%, preferably at least 85%, more preferably at least 90%, still more preferably at least 95%, yet still more preferably at least 99%.

Substrate: a substrate is the molecule that the enzyme naturally recognizes and converts to a product in the biochemical pathway in which the enzyme naturally carries out its function, or is a modified version of the molecule, which is also recognized by the enzyme and is converted by the enzyme to a product in an enzymatic reaction similar to the naturally-occurring reaction.

Tolerance: the ability to continue normal growth or function when exposed to an inhibitor or herbicide in an amount sufficient to suppress the normal growth or function of native, unmodified plants.

Transformation: a process for introducing heterologous DNA into a cell, tissue, or plant. Transformed cells, tissues, or plants are understood to encompass not only the end product of a transformation process, but also transgenic progeny thereof.

Transgenic: stably transformed with a recombinant DNA molecule that preferably comprises a suitable promoter operatively linked to a DNA sequence of interest.

BRIEF DESCRIPTION OF THE SEQUENCES IN THE SEQUENCE LISTING

SEQ ID NO:1 genomic DNA sequence for the *Arabidopsis* 4788 gene

SEQ ID NO:2 cDNA sequence for the *Arabidopsis* 4788 gene

SEQ ID NO:3 amino acid sequence of the *Arabidopsis* 4788 protein

SEQ ID NO:4 oligonucleotide SLP346for

It is an object of the invention to provide an effective and beneficial method to identify novel herbicides. A feature of the invention is the identification of a putative permease gene in *Arabidopsis*. Another feature of the invention is the discovery that the putative permease gene is essential for seedling growth and development. An advantage of the present invention is that the newly discovered essential gene containing a novel herbicidal mode of action enables one skilled in the art to easily and rapidly identify novel herbicides.

One object of the present invention is to provide an essential gene in plants for assay development for inhibitory compounds with herbicidal activity. Genetic results show that when the putative permease gene is mutated in *Arabidopsis*, the resulting phenotype is seedling lethal in the homozygous state. This suggests a critical role for the gene product encoded by the mutated gene.

Using T-DNA insertion mutagenesis, the inventors of the present invention have demonstrated that the activity is essential in *Arabidopsis* seedlings. This implies that chemicals which inhibit the function of the protein in plants are likely to have detrimental effects on plants and are potentially good herbicide candidates. The present invention therefore provides methods of using a purified protein encoded by the gene sequence described below to identify inhibitors thereof, which can then be used as herbicides to suppress the growth of undesirable vegetation, e.g. in fields where crops are grown, particularly agronomically important crops such as maize and other cereal crops such as wheat, oats, rye, sorghum, rice, barley, millet, turf and forage grasses, and the like, as well as cotton, sugar cane, sugar beet, oilseed rape, and soybeans.

The present invention discloses a novel nucleotide sequence derived from *Arabidopsis,* designated the 4788 gene. The nucleotide sequence of the genomic clone is set forth in SEQ ID NO:1, the nucleotide sequence of the corresponding cDNA clone is set forth in SEQ ID NO:2, and the amino acid sequence of the *Arabidopsis* 4788 protein is set forth in SEQ ID NO:3. The present invention also includes nucleotide sequences substantially similar to those set forth in SEQ ID NO:1 and SEQ ID NO: 2.

The present invention also encompasses nucleotide sequences substantially similar to those set forth in SEQ ID NO:1 and SEQ ID NO: 2, wherein said nucleotide sequence is a plant nucleotide sequence. Preferred is a nucleotide sequences substantially similar to those set forth in SEQ ID NO:1 and SEQ ID NO: 2, wherein said nucleotide sequence is an *Arabidopsis thaliana* nucleotide sequence.

Further encompassed is a nucleotide sequence substantially similar to those set forth in SEQ ID NO:1 and SEQ ID NO: 2, wherein the encoded protein has permease activity. More

preferred is a nucleotide sequence substantially similar to those set forth in SEQ ID NO:1 and SEQ ID NO: 2, wherein the encoded protein has purine or pyrimidine permease activity. Particularly preferred is a nucleotide sequence substantially similar to those set forth in SEQ ID NO:1 and SEQ ID NO: 2, wherein said encoded protein has uracil permease activity. Further encompassed is an amino acid sequence comprising an amino acid sequence encoded by a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO: 2. Also encomassed is an amino acid sequence comprising an amino acid sequence encoded by SEQ ID NO: 1 or SEQ ID NO: 2.

The present invention also encompasses proteins whose amino acid sequence are substantially similar to the amino acid sequences set forth in SEQ ID NO:3.

Also encompassed is an amino acid sequence comprising and amino acid sequence substantially similar to SEQ ID NO:3. Preferred is an amino acid sequence comprising and amino acid sequence which is SEQ ID NO:3.

Encompassed is an amino acid sequence comprising amino acid sequence encoded by a nucleotide sequence substantially similar to SEQ ID NO:1 or SEQ ID NO:2., wherein the protein has permease activity. More preferred is an amino acid sequence comprising amino acid sequence encoded by a nucleotide sequence substantially similar to SEQ ID NO:1 or SEQ ID NO:2, wherein the protein has purine or pyrimidine permease activity. Particularly preferred is an amino acid sequence comprising amino acid sequence encoded by a nucleotide sequence substantially similar to SEQ ID NO:1 or SEQ ID NO:2., wherein the protein has uracil permease.

Further encompassed is an amino acid sequence comprising at least 20 consecutive amino acid residues of the amino acid sequence encoded by SEQ ID NO: 1 or SEQ ID NO: 2.

Further encompassed is an amino acid sequence comprising at least 20 consecutive amino acid residues of the amino acid sequence of SEQ ID NO: 3.

An further embodiment is a chimeric gene comprising a promoter operatively linked to a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO:2.

Further encompassed is a recombinant vector comprising a chimeric gene comprising a promoter operatively linked to a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO:2, wherein said vector is capable of being stably transformed into a host cell.

Further encompassed is a host cell comprising a vector comprising a chimeric gene comprising a promoter operatively linked to a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO:2, wherein said vector is capable of being stably transformed into a host cell and wherein said nucleotide sequence is expressible in said cell.

A preferred host cell according to the invention is an eukaryotic cell, more preferred is a host cell selected from the group consisting of an insect cell, a yeast cell, and a plant cell. A further preferred host cell is a prokaryotic cell, more preferred is a bacterial cell.

Also encompassed is a plant comprising a vector comprising a chimeric gene comprising a promoter operatively linked to a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO:2, wherein said vector is capable of being stably transformed into a plant cell; also including the progeny and seed for such a plant, which seed is optionally treated (e.g., primed or coated) and/or packaged, e.g. placed in a bag or other container with instructions for use. More preferred is a plant according to the invention wherin said plant being tolerant to an inhibitor of permease activity; also including the progeny and seed for such a plant, which seed is optionally treated (e.g., primed or coated) and/or packaged, e.g. placed in a bag or other container with instructions for use.

A further embodiment of the invention is a process for making nucleotides sequences encoding gene products having altered permease activity comprising,

a) shuffling a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO:2,

b) expressing the resulting shuffled nucleotide sequences and

c) selecting for altered permease activity as compared to the permease activity of the gene product of said unmodified nucleotide sequence.

Preferred is a process according to the invention, wherein the nucleotide sequence is SEQ ID NO: 1 or SEQ ID NO: 2. More preferred is a process according to the invention, wherein said permease activity is purine or pyrimidine permease activity. Particularly preferred is a

process according to the invention, wherein said permease activity is uracil permease activity.

Further encompassed by the invention is a shuffled DNA molecule obtainable by the process for making nucleotides sequences encoding gene products having altered permease activity comprising,

a) shuffling a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO:2,

b) expressing the resulting shuffled nucleotide sequences and

c) selecting for altered permease activity as compared to the permease activity of the gene product of said unmodified nucleotide sequence.

A further embodiment of the invention is a shuffled DNA molecule produced by a process according to the invention.

Also comprised by the invention is a shuffled DNA molecule obtained a process according to the invention, wherein said shuffled DNA molecule encodes a gene product having enhanced tolerance to an inhibitor of permease activity.

A further embodiment of the invention is a chimeric gene comprising a promoter operatively linked to a shuffled DNA molecule produced by a process according to the invention.

A further embodiment of the invention is a recombinant vector comprising a chimeric gene comprising a promoter operatively linked to a shuffled DNA molecule produced by a process according to the invention, wherein said vector is capable of being stably transformed into a host cell.

Further encompassed is a host cell comprising a vector according to claim comprising a chimeric gene comprising a promoter operatively linked to a shuffled DNA molecule produced by said process, wherein said vector is capable of being stably transformed into a host cell and wherein said nucleotide sequence is expressible in said cell. Preferred is a host cell according to the invention which is an eukaryotic cell, more preferred wherein said host cell is selected from the group consisting of an insect cell, a yeast cell, and a plant cell.

Also preferred is a host cell according to the invention which is a prokaryotic cell, more preferred is a host cell according to the invention which is a bacterial cell.

A further embodiment is a plant comprising a plant cell comprising a vector according according to the invention comprising a chimeric gene comprising a promoter operatively linked to a shuffled DNA molecule produced by the process according to the invention, wherein said vector is capable of being stably transformed into a plant cell and wherein said nucleotide sequence is expressible in said cell; also including the progeny and seed for such a plant, which seed is optionally treated (e.g., primed or coated) and/or packaged, e.g. placed in a bag or other container with instructions for use.

A further embodiment is a plant according to the invention which is tolerant to an inhibitor of permease activity; also including the progeny and seed for such a plant, which seed is optionally treated (e.g., primed or coated) and/or packaged, e.g. placed in a bag or other container with instructions for use.

Further embodied is a process of identifying compounds having herbicidal activity comprising:
a) combining a protein comprising an amino acid sequence encoded by a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO: 2.
and a compound to be tested for the ability to bind to said protein, under conditions conducive to binding,

b) selecting a compound identified in step (a) that is capable of binding said protein,

c) applying identified compound in step (b) to a plant to test for herbicidal activity,

and d) selecting compounds having herbicidal activity.

Further encompassed is a compound having herbicidal activity identifiable according to the process according to the invention.
Further encompassed is a process of identifying an inhibitor of permease activity having herbicidal activity comprising:

a) combining a permease and a compound to be tested for the ability to inhibit the activity of said permease, under conditions conducive to such inhibition,

b) selecting a compound identified in step (a) that is capable of inhibiting said permease activity,

c) applying identified compound in step (b) to a plant to test for herbicidal activity,

and d) selecting compounds having herbicidal activity.

Encompassed by the invention is a process according to the invention, wherein said permease is a purine or pyrimidine permease, more preferred, wherein said permease is a uracil permease.

A further embodiment of the invention is a compound having herbicidal activity identifiable according to the process according to the invnetion.

A further embodiment is a process of identifying an inhibitor of permease activity comprising,

a) introducing SEQ ID NO: 1 or SEQ ID NO:2, or nucleotide sequences substantially similar thereto into a permease-deficient host cell, such as *E. coli uraA* such that said sequence is functionally expressable;

b) combining said host cell containing said nucleotide sequence, with a minimal inhibitory concentration of 5-fluorouracil, and with a compound to be tested for the ability to inhibit the activity of said permease, under conditions conducive to such inhibition,

c) measure host cell growth under the conditions of step (b); and

d) selecting said compound that inhibits host cell growth in step (c).

Further encompassed is a compound having herbicidal activity identifiable according to the process according to the invention.

Further encompassed is a method for suppressing the growth of a plant comprising, applying to said plant a compound that inhibits the activity of the amino acid sequence comprising an amino acid sequence encoded by a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO: 2 in an amount sufficient to suppress the growth of said plant.

Further encompassed is a process of identifying compounds having herbicidal activity comprising:

a) combining a protein of claim 10 and a compound to be tested for the ability to bind to said protein, under conditions conducive to binding,

b) selecting a compound identified in step (a) that is capable of binding said protein,

c) applying identified compound in step (b) to a plant to test for herbicidal activity,

and d) selecting compounds having herbicidal activity.

and the compounds having herbicidal activity identifiable according to the process according to the invention.

Further encompassed is a method of improving crops comprising, applying to a herbicide tolerant plant or seed selected from the group consisting of the plant or seed according to the invention wherein said plant is tolerant to an inhibitor of permease, a compound according to the invention having herbicidal activity in an amount that inhibits the growth of undesired vegetation without significantly suppressing the growth of the herbicide tolerant plant or seed.

The present invention also includes methods of using the 4788 gene product as an herbicide target, based on the essentiality of the gene for normal growth and development. Furthermore, the invention can be used in a screening assay to identify inhibitors that are potential herbicides.

In another preferred embodiment, the present invention describes a method for identifying chemicals having the ability to inhibit 4788 activity in plants preferably comprising the steps of: a) obtaining transgenic plants, plant tissue, plant seeds or plant cells, preferably stably transformed, comprising a non-native nucleotide sequence encoding an enzyme having 4788 activity and capable of overexpressing an enzymatically active 4788 gene product; b) applying a chemical to the transgenic plants, plant cells, tissues or parts and to the isogenic non-transformed plants, plant cells, tissues or parts; c) determining the growth or viability of the transgenic and non-transformed plants, plant cells, tissues after application of the chemical; d) comparing the growth or viability of the transgenic and non-transformed plants, plant cells, tissues after application of the chemical; and e) selecting chemicals suppress the viability or growth of the non-transgenic plants, plant cells, tissues

or parts, without significantly suppressing the growth of the viability or growth of the isogenic transgenic plants, plant cells, tissues or parts. In a preferred embodiment, the enzyme having 4788 activity is encoded by a nucleotide sequence derived from a plant, preferably *Arabidopsis thaliana*, desirably identical or substantially similar to the nucleotide sequence set forth in SEQ ID NO:1 or SEQ ID NO:2. In another embodiment, the enzyme having 4788 activity is encoded by a nucleotide sequence capable of encoding the amino acid sequence of SEQ ID NO:3. In yet another embodiment, the enzyme having 4788 activity has an amino acid sequence identical or substantially similar to the amino acid sequence set forth in SEQ ID NO:3.

The present invention further embodies plants, plant tissues, and plant cells that have modified 4788 activity and that are therefore tolerant to inhibition by a herbicide at levels normally inhibitory to naturally occurring 4788 activity; also including the progeny and seed for such a plant, which seed is optionally treated (e.g., primed or coated) and/or packaged, e.g. placed in a bag or other container with instructions for use. Herbicide tolerant plants encompassed by the invention include those that would otherwise be potential targets for normally inhibiting herbicides, particularly the agronomically important crops mentioned above. According to this embodiment, plants, plant tissue, plant seeds, or plant cells are transformed, preferably stably transformed, with a recombinant DNA molecule comprising a suitable promoter functional in plants operatively linked to a nucleotide coding sequence that encodes a modified 4788 gene that is tolerant to inhibition by a herbicide at a concentration that would normally inhibit the activity of wild-type, unmodified 4788 gene product. Modified 4788 activity may also be conferred upon a plant by increasing expression of wild-type herbicide-sensitive 4788 protein by providing multiple copies of wild-type 4788 genes to the plant or by overexpression of wild-type 4788 genes under control of a stronger-than-wild-type promoter. The transgenic plants, plant tissue, plant seeds, or plant cells thus created are then selected by conventional selection techniques, whereby herbicide tolerant lines are isolated, characterized, and developed. Alternately, random or site-specific mutagenesis may be used to generate herbicide tolerant lines.

Therefore, the present invention provides a plant, plant cell, or plant tissue transformed with a DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme having 4788 activity, also including the progeny and seed for such a plant, which seed is optionally treated (e.g., primed or coated) and/or packaged, e.g. placed in a bag or other container with instructions for use, wherein the enzyme has 4788

activity and wherein the DNA molecule confers upon the plant, plant cell, plant seed, or plant tissue tolerance to a herbicide in amounts that normally inhibits naturally occurring 4788 activity. According to one example of this embodiment, the enzyme having 4788 activity is encoded by a nucleotide sequence identical or substantially similar to the nucleotide sequence set forth in SEQ ID NO:1 or SEQ ID NO:2, or has an amino acid sequence identical or substantially similar to the amino acid sequence set forth in SEQ ID NO:3.

The invention also provides a method for suppressing the growth of a plant comprising the step of applying to the plant a chemical that inhibits the naturally occurring 4788 activity in the plant. In a related aspect, the present invention is directed to a method for selectively suppressing the growth of undesired vegetation in a field containing a crop of planted crop seeds or plants, comprising the steps of: (a) optionally planting herbicide tolerant crops or crop seeds, which are plants or plant seeds that are tolerant to a herbicide that inhibits the naturally occurring 4788 activity; and (b) applying to the crops or crop seeds and the undesired vegetation in the field a herbicide in amounts that inhibit naturally occurring 4788 activity, wherein the herbicide suppresses the growth of the weeds without significantly suppressing the growth of the crops.

Other objects and advantages of the present invention will become apparent to those skilled in the art from a study of the following description of the invention and non-limiting examples.

I. Essentiality of the 4788 Gene in *Arabidopsis* Demonstrated by T-DNA Insertion Mutagenesis

As shown in the examples below, the identification of a novel gene structure, as well as the essentiality of the 4788 gene for normal plant growth and development, have been demonstrated for the first time in *Arabidopsis* using T-DNA insertion mutagenesis. Having established the essentiality of 4788 function in plants and having identified the gene encoding this essential activity, the inventors thereby provide an important and sought after tool for new herbicide development.

Arabidopsis insertional mutant lines segregating for seedling lethal mutations are identified as a first step in the identification of essential proteins. Starting with T2 seeds collected

from single T1 plants containing T-DNA insertions in their genomes, those lines segregating homozygous seedling lethal seedlings are identified. These lines are found by placing seeds onto minimal plant growth media, which contain the fungicides benomyl and maxim, and screening for inviable seedlings after 7 and 14 days in the light at room temperature. Inviable phenotypes include altered pigmentation or altered morphology. These phenotypes are observed either on plates directly or in soil following transplantation of seedlings.

When a line is identified as segregating a seedling lethal, it is determined if the resistance marker in the T-DNA co-segregates with the lethality (Errampalli et al. (1991) The Plant Cell, 3:149-157). Co-segregation analysis is done by placing the seeds on media containing the selective agent and scoring the seedlings for resistance or sensitivity to the agent. Examples of selective agents used are hygromycin or phosphinothricin. About 35 resistant seedlings are transplanted to soil and their progeny are examined for the segregation of the seedling lethal. In the case in which the T-DNA insertion disrupts an essential gene, there is co-segregation of the resistance phenotype and the seedling lethal phenotype in every plant. Therefore, in such a case, all resistant plants segregate seedling lethals in the next generation; this result indicates that each of the resistant plants are heterozygous for the DNA causing both phenotypes.

For those lines showing co-segregation of the T-DNA resistance marker and the seedling lethal phenotype, Southern analysis is performed as an initial step in the characterization of the molecular nature of each insertion. Southerns are done with genomic DNA isolated from heterozygotes and using probes capable of hybridizing with the T-DNA vector DNA. Often, the T-DNA insertion in a given plant is shown to contain multiple copies of the T-DNA vector that insert at a single genetic locus. Using the results of the Southern, appropriate restriction enzymes are chosen to perform plasmid rescue in order to molecularly clone *Arabidopsis* genomic DNA flanking one or both sides of the T-DNA insertion. Plasmids obtained in this manner are analyzed by restriction enzyme digestion to sort the plasmids into classes based on their digestion pattern. For each class of plasmid clone, the DNA sequence is determined. The resulting sequences are analyzed for the presence of non-T-DNA vector sequences. When such sequences are found, they are used to search DNA and protein databases using the BLAST and BLAST2 programs (Altschul et al. (1990) J Mol. Biol. 215: 403-410; Altschul et al (1997) Nucleic Acid Res. 25:3389-3402). Additional genomic and cDNA sequences for each gene are identified by standard molecular biology procedures.

II. Sequence of the *Arabidopsis* 4788 Gene

The *Arabidopsis* 4788 gene is sequenced by isolating DNA flanking the T-DNA border from the tagged seedling-lethal line #4788. *Arabidopsis* DNA flanking the T-DNA border is identical to a internal region of a sequenced BAC of *Arabidopsis* (BAC T9J22, chromosome 2). This BAC clone contains 115,851 bp of sequence, of which a very small portion corresponds to the genomic region that contains the 4788 gene. Nothwithstanding the BAC information, the inventors are the first to establish definitively the entire gene sequence, and to demonstrate for the first time that the 4788 gene product is essential for normal growth and development, as well as defining the function of the 4788 gene product. The present invention discloses the nucleotide sequence of the *Arabidopsis* 4788 gene as well as the amino acid sequence of the *Arabidopsis* 4788 protein. The nucleotide sequence corresponding to the genomic clone is set forth in SEQ ID NO:1, the corresponding cDNA clone is set forth in SEQ ID NO:2, and the amino acid sequence encoding the mature protein is set forth in SEQ ID NO:3. The present invention also encompasses an isolated amino acid sequence derived from a plant, wherein said amino acid sequence is identical or substantially similar to the amino acid sequence encoded by the nucleotide sequence set forth in SEQ ID NO: 1 or SEQ ID NO:2, wherein said amino acid sequence has 4788 activity. Using BLAST and BLAST2 programs with the default settings, the sequence of the 4788 gene shows similarity to uracil permeases.

The 4788 gene is also a member of a gene family in *Arabidopsis*. This gene family consists of at least six members (Genbank Acc. #s: AC002505 (2739376), BAC T9J22; AC004481 (3337350), BAC F13P17; AC001229 (2190545), BAC F5I14; AB009053 (n/a), clone MQB2 (13th ORF); U83501 (1791307); and AA712474 (EST clone 194H6T7) as well as AA605567 (EST clone 205J16XP).

III. Recombinant Production of 4788 and Uses Thereof

For recombinant production of 4788 in a host organism, a nucleotide sequence encoding 4788 protein is inserted into an expression cassette designed for the chosen host and introduced into the host where it is recombinantly produced. The choice of specific regulatory sequences such as promoter, signal sequence, 5' and 3' untranslated sequences, and enhancer appropriate for the chosen host is within the level of skill of the

routineer in the art. The resultant molecule, containing the individual elements operably linked in proper reading frame, may be inserted into a vector capable of being transformed into the host cell. Suitable expression vectors and methods for recombinant production of proteins are well known for host organisms such as *E. coli,* yeast, and insect cells (see, *e.g.*, Luckow and Summers, *Bio/Technol. 6:* 47 (1988), and baculovirus expression vectors, e.g., those derived from the genome of *Autographica californica* nuclear polyhedrosis virus (AcMNPV). A preferred baculovirus/insect system is pAcHLT (Pharmingen, San Diego, CA) used to transfect *Spodoptera frugiperda* Sf9 cells (ATCC) in the presence of linear *Autographa californica* baculovirus DNA (Pharmigen, San Diego, CA). The resulting virus is used to infect HighFive *Tricoplusia ni* cells (Invitrogen, La Jolla, CA).

In a preferred embodiment, the nucleotide sequence encoding a protein having 4788 activity is derived from an eukaryote, such as a mammal, a fly or a yeast, but is preferably derived from a plant. In a further preferred embodiment, the nucleotide sequence is identical or substantially similar to the nucleotide sequence set forth in SEQ ID NO:1 or SEQ ID NO:2, or encodes a protein having 4788 activity, whose amino acid sequence is identical or substantially similar to the amino acid sequence set forth in SEQ ID NO:3. The nucleotide sequence set forth in SEQ ID NO:2 encodes the *Arabidopsis* 4788 protein, whose amino acid sequence is set forth in SEQ ID NO:3. In another preferred embodiment, the nucleotide sequence is derived from a prokaryote, preferably a bacteria, e.g. the uraA gene in *E. coli* (Andersen et al. (1995) J. Bacteriol. 177: 2008-2013). Recombinantly produced 4788 is isolated and purified using a variety of standard techniques. The actual techniques that may be used will vary depending upon the host organism used, whether the protein is designed for secretion, and other such factors familiar to the skilled artisan (*see, e.g.* chapter 16 of Ausubel, F. *et al.*, "Current Protocols in Molecular Biology", pub. by John Wiley & Sons, Inc. (1994).

Assays for Characterizing the 4788 Protein

Recombinantly produced 4788 proteins are useful for a variety of purposes. For example, they can be used in *in vitro* assays to screen known herbicidal chemicals whose target has not been identified to determine if they inhibit 4788. Such *in vitro* assays may also be used as more general screens to identify chemicals that inhibit such enzymatic

activity and that are therefore novel herbicide candidates. Alternatively, recombinantly produced 4788 proteins may be used to elucidate the complex structure of these molecules and to further characterize their association with known inhibitors in order to rationally design new inhibitory herbicides as well as herbicide tolerant forms of the enzymes. Nucleotide sequences substantially similar to SEQ ID NO:1 or SEQ ID NO:2 and proteins substantially similar to SEQ ID NO:3 from any source, including microbial sources, can be used in the assays exemplified herein. Desirably such nucleotide sequences and proteins are derived from plants. More desirably, they are derived from dicot plants. Alternatively, such nucleotide sequences and proteins are derived from non-maize sources, alternatively from non-monocot sources.

Assay for Permease Activity

The 4788 gene product is believed to function as a permease, more specifically as a uracil permease similar to the maize Lpe1 protein (Schultes et al. (1996) The Plant Cell, 8: 463-475). The *leaf permease1* gene product of maize has similarity to purine and pyrimidine permeases from bacteria and fungi, and mutations in *lpe1* are known to adversely affect chloroplast development. Therefore, the *lpe1*-encoded protein is a potential herbicide target, which is further supported by the demonstration herein of the essentiality of the *Arabidopsis* 4788 gene product. A novel assay can be developed based on expression of the plant protein in a bacterial host lacking the corresponding activity (Andersen et al. (1995) J. Bacteriol. 177: 2008-2013).

A simple assay can be developed to screen for compounds that affect normal functioning of the plant-encoded activity. Such compounds are promising *in vitro* leads that can be tested for *in vivo* herbicidal activity. One assay consists of growing *E. coli uraA* harboring and functionally expressing the 4788 gene in minimal medium in the presence of a minimal inhibitory concentration of 5-fluorouracil. This is accomplished in a 96-well format for automated high-throughput screening. Compounds that are effective in blocking function of the 4788 protein inhibits the ability of the cells to take up 5-FU, and bacterial growth results. This growth is measured by simple turbidometric means.

Other assays based on expression of plant genes in corresponding bacterial mutants have been described. However, in addition to being a novel herbicide target, a particular advantage of this assay is that because uracil permease is expressed on the cell surface, compounds that are effective in inhibiting its function need not penetrate the cell in

order to exhibit their activity. This can be a major problem with using bacterial systems as models for finding compounds that will be effective in plants, because many potent herbicides are known to lack activity on bacteria due to poor uptake of the compound by the bacterium.

In Vitro Inhibitor Assays: Discovery of Small Molecule Ligand that Interacts with Protein of Unknown Function

Once a protein has been identified as a potential herbicide target, the next step is to develop an assay that allows screening large number of chemicals to determine which ones interact with the protein. Although it is straightforward to develop assays for proteins of known function, developing assays with proteins of unknown functions is more difficult.

To address this issue, novel technologies are being examined that can detect interactions between a protein and a ligand without knowing the biological function of the protein. A short description of three methods is presented, including fluorescence correlation spectroscopy, surface-enhanced laser desorption/ionization, and biacore technologies. Many more of these methods are currently being discovered, and some may be amenable to automated, large scale screening in light of this disclosure.

Fluorescence Correlation Spectroscopy (FCS) theory was developed in 1972 but it is only in recent years that the technology to perform FCS became available (Madge et al. (1972) Phys. Rev. Lett., 29: 705-708; Maiti et al. (1997) Proc. Natl. Acad. Sci. USA, 94: 11753-11757). FCS measures the average diffusion rate of a fluorescent molecule within a small sample volume. The sample size can be as low as 10^3 fluorescent molecules and the sample volume as low as a the cytoplasm of a single bacterium. The diffusion rate is a function of the mass of the molecule and decreases as the mass increases. FCS can therefore be applied to protein-ligand interaction analysis by measuring the change in mass and therefore in diffusion rate of a molecule upon binding.

Surface-Enhanced Laser Desorption/Ionization (SELDI) was invented by Hutchens and Yip during the late 1980's (Hutchens and Yip (1993) Rapid Commun. Mass Spectrom. 7: 576-580). When coupled to a time-of-flight mass spectrometer (TOF), SELDI provides a mean to rapidly analyze molecules retained on a chip. It can be applied to ligand-protein interaction analysis by covalently binding the target protein on the chip and analyze by MS the small molecules retained by this protein (Worrall et al. (1998) Anal. Biochem. 70: 750-756).

Biacore relies on changes in the refractive index at the surface layer upon binding of a ligand to a protein immobilized on the layer. In this system, a collection of small ligands is injected sequentially in a 2-5 ul cell with the immobilized protein. Binding is detected by surface plasmon resonance (SPR) by recording laser light refracting from the surface. In general, the refractive index change for a given change of mass concentration at the surface layer, is practically the same for all proteins and peptides, allowing a single method to be applicable for any protein (Liedberg et al. (1983) Sensors Actuators 4: 299-304; Malmquist (1993) Nature, 361: 186-187).

IV. *In Vivo* Inhibitor Assay

In one embodiment, a suspected herbicide, for example identified by *in vitro* screening, is applied to plants at various concentrations. The suspected herbicide is preferably sprayed on the plants. After application of the suspected herbicide, its effect on the plants, for example death or suppression of growth is recorded.

In another embodiment, an *in vivo* screening assay for inhibitors of the 4788 activity uses transgenic plants, plant tissue, plant seeds or plant cells capable of overexpressing a nucleotide sequence having 4788 activity, wherein the 4788 gene product is enzymatically active in the transgenic plants, plant tissue, plant seeds or plant cells. The nucleotide sequence is preferably derived from an eukaryote, such as a yeast, but is preferably derived from a plant. In a further preferred embodiment, the nucleotide sequence is identical or substantially similar to the nucleotide sequence set forth in SEQ ID NO:1 or SEQ ID NO:2, or encodes an enzyme having 4788 activity, whose amino acid sequence is identical or substantially similar to the amino acid sequence set forth in SEQ ID NO:3. In another preferred embodiment, the nucleotide sequence is derived from a prokaryote, preferably a bacteria, e.g. the uraA gene of *E. coli*.

A chemical is then applied to the transgenic plants, plant tissue, plant seeds or plant cells and to the isogenic non-transgenic plants, plant tissue, plant seeds or plant cells, and the growth or viability of the transgenic and non-transformed plants, plant tissue, plant seeds or plant cells are determined after application of the chemical and compared. Compounds capable of inhibiting the growth of the non-transgenic plants, but not affecting the growth of the transgenic plants are selected as specific inhibitors of 4788 activity.

V.Herbicide Tolerant Plants

The present invention is further directed to plants, plant tissue, plant seeds, and plant cells tolerant to herbicides that inhibit the naturally occurring 4788 activity in these plants, wherein the tolerance is conferred by an altered 4788 activity. Altered 4788 activity may be conferred upon a plant according to the invention by increasing expression of wild-type herbicide-sensitive 4788 by providing additional wild-type 4788 genes to the plant, by expressing modified herbicide-tolerant 4788 genes in the plant, or by a combination of these techniques. Representative plants include any plants to which these herbicides are applied for their normally intended purpose. Preferred are agronomically important crops such as cotton, soybean, oilseed rape, sugar beet, maize, rice, wheat, barley, oats, rye, sorghum, millet, turf, forage, turf grasses, and the like.

A. Increased Expression of Wild-Type 4788

Achieving altered 4788 activity through increased expression results in a level of a 4788 in the plant cell at least sufficient to overcome growth inhibition caused by the herbicide. The level of expressed enzyme generally is at least two times, preferably at least five times, and more preferably at least ten times the natively expressed amount. Increased expression may be due to multiple copies of a wild-type 4788 gene; multiple occurrences of the coding sequence within the gene (*i.e.* gene amplification) or a mutation in the non-coding, regulatory sequence of the endogenous gene in the plant cell. Plants having such altered gene activity can be obtained by direct selection in plants by methods known in the art (see, *e.g.* U.S. Patent No. 5,162,602, and U.S. Patent No. 4,761,373, and references cited therein). These plants also may be obtained by genetic engineering techniques known in the art. Increased expression of a herbicide-sensitive 4788 gene can also be accomplished by transforming a plant cell with a recombinant or chimeric DNA molecule comprising a promoter capable of driving expression of an associated structural gene in a plant cell operatively linked to a homologous or heterologous structural gene encoding the 4788 protein. Preferably, the transformation is stable, thereby providing a heritable transgenic trait.

B. Expression of Modified Herbicide-Tolerant 4788 Proteins

According to this embodiment, plants, plant tissue, plant seeds, or plant cells are stably transformed with a recombinant DNA molecule comprising a suitable promoter functional in plants operatively linked to a coding sequence encoding a herbicide tolerant form of 4788. A herbicide tolerant form of the enzyme has at least one amino acid substitution, addition or deletion that confers tolerance to a herbicide that inhibits the unmodified, naturally occurring form of the enzyme. The transgenic plants, plant tissue, plant seeds, or plant cells thus created are then selected by conventional selection techniques, whereby herbicide tolerant lines are isolated, characterized, and developed. Below are described methods for obtaining genes that encode herbicide tolerant forms of 4788:

One general strategy involves direct or indirect mutagenesis procedures on microbes. For instance, a genetically manipulatable microbe such as *E. coli* or *S. cerevisiae* may be subjected to random mutagenesis *in vivo* with mutagens such as UV light or ethyl or methyl methane sulfonate. Mutagenesis procedures are described, for example, in Miller, *Experiments in Molecular Genetics,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1972); Davis *et al.*, *Advanced Bacterial Genetics,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1980); Sherman *et al., Methods in Yeast Genetics,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1983); and U.S. Patent No. 4,975,374. The microbe selected for mutagenesis contains a normal, inhibitor-sensitive 4788 gene and is dependent upon the activity conferred by this gene. The mutagenized cells are grown in the presence of the inhibitor at concentrations that inhibit the unmodified gene. Colonies of the mutagenized microbe that grow better than the unmutagenized microbe in the presence of the inhibitor (i.e. exhibit resistance to the inhibitor) are selected for further analysis. 4788 genes from these colonies are isolated, either by cloning or by PCR amplification, and their sequences are elucidated. Sequences encoding altered gene products are then cloned back into the microbe to confirm their ability to confer inhibitor tolerance.

A method of obtaining mutant herbicide-tolerant alleles of a plant 4788 gene involves direct selection in plants. For example, the effect of a mutagenized 4788 gene on the growth inhibition of plants such as *Arabidopsis*, soybean, or maize is determined by plating seeds sterilized by art-recognized methods on plates on a simple minimal salts medium containing increasing concentrations of the inhibitor. Such concentrations are in the range of 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 110, 300, 1000 and 3000 parts

per million (ppm). The lowest dose at which significant growth inhibition can be reproducibly detected is used for subsequent experiments. Determination of the lowest dose is routine in the art.

Mutagenesis of plant material is utilized to increase the frequency at which resistant alleles occur in the selected population. Mutagenized seed material is derived from a variety of sources, including chemical or physical mutagenesis or seeds, or chemical or physical mutagenesis or pollen (Neuffer, In *Maize for Biological Research* Sheridan, ed. Univ. Press, Grand Forks, ND., pp. 61-64 (1982)), which is then used to fertilize plants and the resulting M_1 mutant seeds collected. Typically for *Arabidopsis*, M_2 seeds (Lehle Seeds, Tucson, AZ), which are progeny seeds of plants grown from seeds mutagenized with chemicals, such as ethyl methane sulfonate, or with physical agents, such as gamma rays or fast neutrons, are plated at densities of up to 10,000 seeds/plate (10 cm diameter) on minimal salts medium containing an appropriate concentration of inhibitor to select for tolerance. Seedlings that continue to grow and remain green 7-21 days after plating are transplanted to soil and grown to maturity and seed set. Progeny of these seeds are tested for tolerance to the herbicide. If the tolerance trait is dominant, plants whose seed segregate 3:1 / resistant:sensitive are presumed to have been heterozygous for the resistance at the M_2 generation. Plants that give rise to all resistant seed are presumed to have been homozygous for the resistance at the M_2 generation. Such mutagenesis on intact seeds and screening of their M2 progeny seed can also be carried out on other species, for instance soybean (*see, e.g.* U.S. Pat. No. 5,084,082). Alternatively, mutant seeds to be screened for herbicide tolerance are obtained as a result of fertilization with pollen mutagenized by chemical or physical means.

Confirmation that the genetic basis of the herbicide tolerance is a 4788 gene is ascertained as exemplified below. First, alleles of the 4788 gene from plants exhibiting resistance to the inhibitor are isolated using PCR with primers based either upon the *Arabidopsis* cDNA coding sequences shown in SEQ ID NO:2 or, more preferably, based upon the unaltered 4788 gene sequence from the plant used to generate tolerant alleles. After sequencing the alleles to determine the presence of mutations in the coding sequence, the alleles are tested for their ability to confer tolerance to the inhibitor on plants into which the putative tolerance-conferring alleles have been transformed. These plants can be either *Arabidopsis* plants or any other plant whose growth is susceptible to the 4788 inhibitors. Second, the inserted 4788 genes are mapped relative to known restriction fragment length polymorphisms (RFLPs) (*See, for example*, Chang *et al. Proc. Natl. Acad,*

Sci, USA 85: 6856-6860 (1988); Nam *et al., Plant Cell* 1: 699-705 (1989), cleaved amplified polymorphic sequences (CAPS) (Konieczny and Ausubel (1993) The Plant Journal, 4(2): 403-410), or SSLPs (Bell and Ecker (1994) Genomics, 19: 137-144). The 4788 inhibitor tolerance trait is independently mapped using the same markers. When tolerance is due to a mutation in that 4788 gene, the tolerance trait maps to a position indistinguishable from the position of the 4788 gene.

Another method of obtaining herbicide-tolerant alleles of a 4788 gene is by selection in plant cell cultures. Explants of plant tissue, *e.g.* embryos, leaf disks, etc. or actively growing callus or suspension cultures of a plant of interest are grown on medium in the presence of increasing concentrations of the inhibitory herbicide or an analogous inhibitor suitable for use in a laboratory environment. Varying degrees of growth are recorded in different cultures. In certain cultures, fast-growing variant colonies arise that continue to grow even in the presence of normally inhibitory concentrations of inhibitor. The frequency with which such faster-growing variants occur can be increased by treatment with a chemical or physical mutagen before exposing the tissues or cells to the inhibitor. Putative tolerance-conferring alleles of the 4788 gene are isolated and tested as described in the foregoing paragraphs. Those alleles identified as conferring herbicide tolerance may then be engineered for optimal expression and transformed into the plant. Alternatively, plants can be regenerated from the tissue or cell cultures containing these alleles.

Still another method involves mutagenesis of wild-type, herbicide sensitive plant 4788 genes in bacteria or yeast, followed by culturing the microbe on medium that contains inhibitory concentrations of the inhibitor and then selecting those colonies that grow in the presence of the inhibitor. More specifically, a plant cDNA, such as the *Arabidopsis* cDNA encoding the 4788 is cloned into a microbe that otherwise lacks the selected gene's activity. The transformed microbe is then subjected to *in vivo* mutagenesis or to *in vitro* mutagenesis by any of several chemical or enzymatic methods known in the art, e.g. sodium bisulfite (Shortle *et al., Methods Enzymol. 100:*457-468 (1983); methoxylamine (Kadonaga *et al., Nucleic Acids Res. 13:*1733-1745 (1985); oligonucleotide-directed saturation mutagenesis (Hutchinson *et al., Proc. Natl. Acad. Sci. USA, 83:*710-714 (1986); or various polymerase misincorporation strategies (see, e.g. Shortle et al., Proc. Natl. Acad. Sci. USA, 79:1588-1592 (1982); Shiraishi *et al., Gene 64:*313-319 (1988); and Leung *et al., Technique 1:*11-15 (1989). Colonies that grow in the presence of normally inhibitory concentrations of inhibitor are picked and purified by repeated restreaking. Their plasmids are purified and tested for the ability to confer tolerance to the inhibitor by retransforming

them into the microbe lacking 4788 gene activity. The DNA sequences of cDNA inserts from plasmids that pass this test are then determined.

Herbicide resistant 4788 proteins are also obtained using methods involving *in vitro* recombination, also called DNA shuffling. By DNA shuffling, mutations, preferably random mutations, are introduced in 4788 genes. DNA shuffling also leads to the recombination and rearrangement of sequences within an 4788 gene or to recombination and exchange of sequences between two or more different of 4788 genes. These methods allows for the production of millions of mutated 4788 genes. The mutated genes, or shuffled genes, are screened for desirable properties, e.g. improved tolerance to herbicides and for mutations that provide broad spectrum tolerance to the different classes of inhibitor chemistry. Such screens are well within the skills of a routineer in the art.

In a preferred embodiment, a mutagenized 4788 gene is formed from at least one template 4788 gene, wherein the template 4788 gene has been cleaved into double-stranded random fragments of a desired size, and comprising the steps of adding to the resultant population of double-stranded random fragments one or more single or double-stranded oligonucleotides, wherein said oligonucleotides comprise an area of identity and an area of heterology to the double-stranded random fragments; denaturing the resultant mixture of double-stranded random fragments and oligonucleotides into single-stranded fragments; incubating the resultant population of single-stranded fragments with a polymerase under conditions which result in the annealing of said single-stranded fragments at said areas of identity to form pairs of annealed fragments, said areas of identity being sufficient for one member of a pair to prime replication of the other, thereby forming a mutagenized double-stranded polynucleotide; and repeating the second and third steps for at least two further cycles, wherein the resultant mixture in the second step of a further cycle includes the mutagenized double-stranded polynucleotide from the third step of the previous cycle, and the further cycle forms a further mutagenized double-stranded polynucleotide, wherein the mutagenized polynucleotide is a mutated 4788 gene having enhanced tolerance to a herbicide which inhibits naturally occurring 4788 activity. In a preferred embodiment, the concentration of a single species of double-stranded random fragment in the population of double-stranded random fragments is less than 1% by weight of the total DNA. In a further preferred embodiment, the template double-stranded polynucleotide comprises at least about 100 species of polynucleotides. In another preferred embodiment, the size of the double-stranded random fragments is from about 5 bp to 5 kb. In a further preferred embodiment, the fourth step of the method comprises

repeating the second and the third steps for at least 10 cycles. Such method is described e.g. in Stemmer et al. (1994) Nature 370: 389-391, in US Patent 5,605,793 and in Crameri et al. (1998) Nature 391: 288-291, Cherry et al. (1999) Nature Biotechnology, 17: 379-384, as well as in WO 97/20078, and these references are incorporated herein by reference.

In another preferred embodiment, any combination of two or more different 4788 genes are mutagenized *in vitro* by a staggered extension process (StEP), as described e.g. in Zhao et al. (1998) Nature Biotechnology 16: 258-261. The two or more 4788 genes are used as template for PCR amplification with the extension cycles of the PCR reaction preferably carried out at a lower temperature than the optimal polymerization temperature of the polymerase. For example, when a thermostable polymerase with an optimal temperature of approximately 72°C is used, the temperature for the extension reaction is desirably below 72°C, more desirably below 65°C, preferably below 60°C, more preferably the temperature for the extension reaction is 55°C. Additionally, the duration of the extension reaction of the PCR cycles is desirably shorter than usually carried out in the art, more desirably it is less than 30 seconds, preferably it is less than 15 seconds, more preferably the duration of the extension reaction is 5 seconds. Only a short DNA fragment is polymerized in each extension reaction, allowing template switch of the extension products between the starting DNA molecules after each cycle of denaturation and annealing, thereby generating diversity among the extension products. The optimal number of cycles in the PCR reaction depends on the length of the 4788 coding regions to be mutagenized but desirably over 40 cycles, more desirably over 60 cycles, preferably over 80 cycles are used. Optimal extension conditions and the optimal number of PCR cycles for every combination of 4788 genes are determined as described in using procedures well-known in the art. The other parameters for the PCR reaction are essentially the same as commonly used in the art. The primers for the amplification reaction are preferably designed to anneal to DNA sequences located outside of the coding sequence of the 4788 genes, e.g. to DNA sequences of a vector comprising the 4788 genes, whereby the different 4788 genes used in the PCR reaction are preferably comprised in separate vectors. The primers desirably anneal to sequences located less than 500 bp away from 4788 coding sequences, preferably less than 200 bp away from the 4788 coding sequences, more preferably less than 120 bp away from the 4788 coding sequences. Preferably, the 4788 coding sequences are surrounded by restriction sites, which are included in the DNA sequence amplified during the PCR reaction, thereby facilitating the cloning of the amplified products into a suitable vector.

In another preferred embodiment, fragments of 4788 genes having cohesive ends are produced as described in WO 98/05765. The cohesive ends are produced by ligating a first oligonucleotide corresponding to a part of a 4788 gene to a second oligonucleotide not present in the gene or corresponding to a part of the gene not adjoining to the part of the gene corresponding to the first oligonucleotide, wherein the second oligonucleotide contains at least one ribonucleotide. A double-stranded DNA is produced using the first oligonucleotide as template and the second oligonucleotide as primer. The ribonucleotide is cleaved and removed. The nucleotide(s) located 5' to the ribonucleotide is also removed, resulting in double-stranded fragments having cohesive ends. Such fragments are randomly reassembled by ligation to obtain novel combinations of gene sequences.

Herbicide resistant proteins are also obtained using methods involving *in situ* modification of a target gene. A technology for targeting and mutating genes *in vivo* can be used, based on self-complementary chimeric oligonucleotides. This approach is being developed for the modification of endogenous genes in a site-specific and genetically inheritable manner (Beetham et al. (1999) Proc. Natl. Acad. Sci. 96: 8774-8778; U.S. Patent No. 5,756,325; U.S. Patent No. 5,871,984; U.S. Patent 5,731,181), and these references are incorporated herein by reference. Furthermore, methods for producing plants exhibiting agronomically desirable traits comprising mutating or modifying genes *in situ*, in a plant cell, are described (WO98/54330), and this reference is incorporated herein by reference. Such modifications can be made via directed mutagenesis techniques such as homologous recombination and selected for based on the resulting herbicide-resistance phenotype (*see, e.g.* Pazkowski *et al., EMBO J. 7:* 4021-4026 (1988), and U.S. Patent No. 5,487,992, particularly columns 18-19 and Example 8), and these references are incorporated herein by reference.

Any 4788 gene or any combination of 4788 genes is used for *in vitro* recombination in the context of the present invention, for example, an 4788 gene derived from a plant, such as, e.g. *Arabidopsis thaliana*, e.g. an 4788 gene set forth in SEQ ID NO:1 or SEQ ID NO:2, an 4788 gene from a bacteria, such as *Bacillus caldolyticus* (Ghim and Neuhard (1994) J. Bacteriol. 176: 3698-3707) or *E. coli* (Andersen et al. (1995) J. Bacteriol. 177: 2008-2013), a 4788 gene from *Zea mays* (Schultes et al. (1996) The Plant Cell, 8: 463-475), and all of which are incorporated herein by reference. Whole 4788 genes or portions thereof are used in the context of the present invention. The library of mutated 4788 genes obtained by the methods described above are cloned into appropriate expression vectors and the resulting vectors are transformed into an appropriate host, for example an algae

like *Chlamydomonas*, a yeast or a bacteria. An appropriate host is preferably a host that otherwise lacks 4788 gene activity, for example *E. coli* uraA mutant (Andersen et al. (1995) J. Bacteriol. 177: 2008-2013). Host cells transformed with the vectors comprising the library of mutated 4788 genes are cultured on medium that contains inhibitory concentrations of the inhibitor and those colonies that grow in the presence of the inhibitor are selected. Colonies that grow in the presence of normally inhibitory concentrations of inhibitor are picked and purified by repeated restreaking. Their plasmids are purified and the DNA sequences of cDNA inserts from plasmids that pass this test are then determined.

An assay for identifying a modified 4788 gene that is tolerant to an inhibitor may be performed in the same manner as the assay to identify inhibitors of the 4788 activity (Inhibitor Assay, above) with the following modifications: First, a mutant 4788 is substituted in one of the reaction mixtures for the wild-type 4788 of the inhibitor assay. Second, an inhibitor of wild-type enzyme is present in both reaction mixtures. Third, mutated activity (activity in the presence of inhibitor and mutated enzyme) and unmutated activity (activity in the presence of inhibitor and wild-type enzyme) are compared to determine whether a significant increase in enzymatic activity is observed in the mutated activity when compared to the unmutated activity. Mutated activity is any measure of activity of the mutated enzyme while in the presence of a suitable substrate and the inhibitor. Unmutated activity is any measure of activity of the wild-type enzyme while in the presence of a suitable substrate and the inhibitor. A significant increase is defined as an increase in enzymatic activity that is larger than the margin of error inherent in the measurement technique, preferably an increase by at least 2-fold or greater of the activity of the wild-type enzyme in the presence of the inhibitor, more preferably an increase by at least 5-fold or greater, most preferably an increase by at least 10-fold or greater.

In addition to being used to create herbicide-tolerant plants, genes encoding herbicide tolerant 4788 can also be used as selectable markers in plant cell transformation methods. For example, plants, plant tissue, plant seeds, or plant cells transformed with a transgene can also be transformed with a gene encoding an altered 4788 capable of being expressed by the plant. The transformed cells are transferred to medium containing an inhibitor of the enzyme in an amount sufficient to inhibit the survivability of plant cells not expressing the modified gene, wherein only the transformed cells will survive. The method is applicable to any plant cell capable of being transformed with a modified 4788-encoding gene, and can be used with any transgene of interest. Expression of the transgene and the

modified gene can be driven by the same promoter functional in plant cells, or by separate promoters.

VI. Plant Transformation Technology

A wild-type or herbicide-tolerant form of the 4788 gene can be incorporated in plant or bacterial cells using conventional recombinant DNA technology. Generally, this involves inserting a DNA molecule encoding the 4788 into an expression system to which the DNA molecule is heterologous (i.e., not normally present) using standard cloning procedures known in the art. The vector contains the necessary elements for the transcription and translation of the inserted protein-coding sequences in a host cell containing the vector. A large number of vector systems known in the art can be used, such as plasmids, bacteriophage viruses and other modified viruses. The components of the expression system may also be modified to increase expression. For example, truncated sequences, nucleotide substitutions or other modifications may be employed. Expression systems known in the art can be used to transform virtually any crop plant cell under suitable conditions. A transgene comprising a wild-type or herbicide-tolerant form of the 4788 gene is preferably stably transformed and integrated into the genome of the host cells. In another preferred embodiment, the transgene comprising a wild-type or herbicide-tolerant form of the 4788 gene located on a self-replicating vector. Examples of self-replicating vectors are viruses, in particular gemini viruses. Transformed cells can be regenerated into whole plants such that the chosen form of the 4788 gene confers herbicide tolerance in the transgenic plants.

A. Requirements for Construction of Plant Expression Cassettes

Gene sequences intended for expression in transgenic plants are first assembled in expression cassettes behind a suitable promoter expressible in plants. The expression cassettes may also comprise any further sequences required or selected for the expression of the transgene. Such sequences include, but are not restricted to, transcription terminators, extraneous sequences to enhance expression such as introns, vital sequences, and sequences intended for the targeting of the gene product to specific organelles and cell compartments. These expression cassettes can then be easily transferred to the plant transformation vectors described *infra*. The following is a description of various components of typical expression cassettes.

1. Promoters

The selection of the promoter used in expression cassettes will determine the spatial and temporal expression pattern of the transgene in the transgenic plant. Selected promoters will express transgenes in specific cell types (such as leaf epidermal cells, mesophyll cells, root cortex cells) or in specific tissues or organs (roots, leaves or flowers, for example) and the selection will reflect the desired location of accumulation of the gene product. Alternatively, the selected promoter may drive expression of the gene under various inducing conditions. Promoters vary in their strength, i.e., ability to promote transcription. Depending upon the host cell system utilized, any one of a number of suitable promoters known in the art can be used. For example, for constitutive expression, the CaMV 35S promoter, the rice actin promoter, or the ubiquitin promoter may be used. For regulatable expression, the chemically inducible PR-1 promoter from tobacco or *Arabidopsis* may be used (*see, e.g.*, U.S. Patent No. 5,689,044).

2. Transcriptional Terminators

A variety of transcriptional terminators are available for use in expression cassettes. These are responsible for the termination of transcription beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators are those that are known to function in plants and include the CaMV 35S terminator, the *tml* terminator, the nopaline synthase terminator and the pea *rbcS* E9 terminator. These can be used in both monocotyledonous and dicotyledonous plants.

3. Sequences for the Enhancement or Regulation of Expression

Numerous sequences have been found to enhance gene expression from within the transcriptional unit and these sequences can be used in conjunction with the genes of this invention to increase their expression in transgenic plants. For example, various intron sequences such as introns of the maize *Adh1* gene have been shown to enhance expression, particularly in monocotyledonous cells. In addition, a number of non-translated leader sequences derived from viruses are also known to enhance expression, and these are particularly effective in dicotyledonous cells.

4. Coding Sequence Optimization

The coding sequence of the selected gene may be genetically engineered by altering the coding sequence for optimal expression in the crop species of interest. Methods for modifying coding sequences to achieve optimal expression in a particular crop species are well known (see, *e.g.* Perlak *et al.*, *Proc. Natl. Acad. Sci. USA 88:* 3324 (1991); and Koziel *et al.*, *Bio/technol. 11:* 194 (1993)).

5. Targeting of the Gene Product Within the Cell

Various mechanisms for targeting gene products are known to exist in plants and the sequences controlling the functioning of these mechanisms have been characterized in some detail. For example, the targeting of gene products to the chloroplast is controlled by a signal sequence found at the amino terminal end of various proteins which is cleaved during chloroplast import to yield the mature protein (*e.g.* Comai *et al.* J. Biol. Chem. 263: 15104-15109 (1988)). Other gene products are localized to other organelles such as the mitochondrion and the peroxisome (*e.g.* Unger *et al.* Plant Molec. Biol. 13: 411-418 (1989)). The cDNAs encoding these products can also be manipulated to effect the targeting of heterologous gene products to these organelles. In addition, sequences have been characterized which cause the targeting of gene products to other cell compartments. Amino terminal sequences are responsible for targeting to the ER, the apoplast, and extracellular secretion from aleurone cells (Koehler & Ho, Plant Cell 2: 769-783 (1990)). Additionally, amino terminal sequences in conjunction with carboxy terminal sequences are responsible for vacuolar targeting of gene products (Shinshi *et al.* Plant Molec. Biol. 14: 357-368 (1990)). By the fusion of the appropriate targeting sequences described above to transgene sequences of interest it is possible to direct the transgene product to any organelle or cell compartment.

B. Construction of Plant Transformation Vectors

Numerous transformation vectors available for plant transformation are known to those of ordinary skill in the plant transformation arts, and the genes pertinent to this invention can be used in conjunction with any such vectors. The selection of vector will depend upon the preferred transformation technique and the target species for transformation. For certain target species, different antibiotic or herbicide selection markers may be preferred. Selection markers used routinely in transformation include the *nptII* gene, which confers resistance to kanamycin and related antibiotics (Messing & Vierra. Gene 19: 259-268

(1982); Bevan et al., Nature 304:184-187 (1983)), the *bar* gene, which confers resistance to the herbicide phosphinothricin (White et al., Nucl. Acids Res 18: 1062 (1990), Spencer et al. Theor. Appl. Genet 79: 625-631 (1990)), the *hph* gene, which confers resistance to the antibiotic hygromycin (Blochinger & Diggelmann, Mol Cell Biol 4: 2929-2931), and the *dhfr* gene, which confers resistance to methotrexate (Bourouis et al., EMBO J. 2(7): 1099-1104 (1983)), and the EPSPS gene, which confers resistance to glyphosate (U.S. Patent Nos. 4,940,935 and 5,188,642).

1. Vectors Suitable for *Agrobacterium* Transformation

Many vectors are available for transformation using *Agrobacterium tumefaciens*. These typically carry at least one T-DNA border sequence and include vectors such as pBIN19 (Bevan, Nucl. Acids Res. (1984)) and pXYZ. Typical vectors suitable for *Agrobacterium* transformation include the binary vectors pCIB200 and pCIB2001, as well as the binary vector pCIB10 and hygromycin selection derivatives thereof. (*See*, for example, U.S. Patent No. 5,639,949).

2. Vectors Suitable for non-*Agrobacterium* Transformation

Transformation without the use of *Agrobacterium tumefaciens* circumvents the requirement for T-DNA sequences in the chosen transformation vector and consequently vectors lacking these sequences can be utilized in addition to vectors such as the ones described above which contain T-DNA sequences. Transformation techniques that do not rely on *Agrobacterium* include transformation via particle bombardment, protoplast uptake (*e.g.* PEG and electroporation) and microinjection. The choice of vector depends largely on the preferred selection for the species being transformed. Typical vectors suitable for non-*Agrobacterium* transformation include pCIB3064, pSOG19, and pSOG35. (*See*, for example, U.S. Patent No. 5,639,949).

C. Transformation Techniques

Once the coding sequence of interest has been cloned into an expression system, it is transformed into a plant cell. Methods for transformation and regeneration of plants are well known in the art. For example, Ti plasmid vectors have been utilized for the delivery of foreign DNA, as well as direct DNA uptake, liposomes, electroporation, micro-injection, and

microprojectiles. In addition, bacteria from the genus *Agrobacterium* can be utilized to transform plant cells.

Transformation techniques for dicotyledons are well known in the art and include *Agrobacterium*-based techniques and techniques that do not require *Agrobacterium*. Non-*Agrobacterium* techniques involve the uptake of exogenous genetic material directly by protoplasts or cells. This can be accomplished by PEG or electroporation mediated uptake, particle bombardment-mediated delivery, or microinjection. In each case the transformed cells are regenerated to whole plants using standard techniques known in the art. Transformation of most monocotyledon species has now also become routine. Preferred techniques include direct gene transfer into protoplasts using PEG or electroporation techniques, particle bombardment into callus tissue, as well as *Agrobacterium*-mediated transformation.

VII. Breeding

The wild-type or altered form of a 4788 gene of the present invention can be utilized to confer herbicide tolerance to a wide variety of plant cells, including those of gymnosperms, monocots, and dicots. Although the gene can be inserted into any plant cell falling within these broad classes, it is particularly useful in crop plant cells, such as rice, wheat, barley, rye, corn, potato, carrot, sweet potato, sugar beet, bean, pea, chicory, lettuce, cabbage, cauliflower, broccoli, turnip, radish, spinach, asparagus, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, quince, melon, plum, cherry, peach, nectarine, apricot, strawberry, grape, raspberry, blackberry, pineapple, avocado, papaya, mango, banana, soybean, tobacco, tomato, sorghum and sugarcane.
The high-level expression of a wild-type 4788 gene and/or the expression of herbicide-tolerant forms of a 4788 gene conferring herbicide tolerance in plants, in combination with other characteristics important for production and quality, can be incorporated into plant lines through breeding approaches and techniques known in the art.
Where a herbicide tolerant 4788 gene allele is obtained by direct selection in a crop plant or plant cell culture from which a crop plant can be regenerated, it is moved into commercial varieties using traditional breeding techniques to develop a herbicide tolerant crop without the need for genetically engineering the allele and transforming it into the plant.

The invention will be further described by reference to the following detailed examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless otherwise specified.

EXAMPLES

Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by Sambrook, *et al.*, Molecular Cloning, eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989) and by T.J. Silhavy, M.L. Berman, and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and by Ausubel, F.M. *et al.*, Current Protocols in Molecular Biology, pub. by Greene Publishing Assoc. and Wiley-Interscience (1987), Reiter, et al., Methods in *Arabidopsis* Research, World Scientific Press (1992), and Schultz et al., Plant Molecular Biology Manual, Kluwer Academic Publishers (1998). These references describe the standard techniques used for all steps in tagging and cloning genes from T-DNA mutagenized populations of *Arabidopsis*: plant infection and transformation; screening for the identification of seedling mutants; cosegregation analysis; and plasmid rescue.

Example 1: Sequence Analysis of Tagged Seedling – Lethal Line #4788 From the T-DNA Mutagenized Population of *Arabidopsis*

The plasmid rescue technique is used to molecularly clone *Arabidopsis* genomic DNA flanking one or both sides of T-DNA insertions resulting from T-DNA mutagenesis. Plasmids obtained in this manner are analyzed by restriction enzyme digestion to sort the plasmids into classes based on their digestion pattern. For each class of plasmid clone, the DNA sequence is determined. The resulting sequences are analyzed for the presence of non-T-DNA vector sequences. The plasmids recovered from the plasmid rescue protocol are sequenced using the slp346 primer. Primer slp346 provides information on the flanking sequence immediately adjacent to the left T-DNA border. Plasmid rescue is validated by PCR of genomic DNA from a heterozygote for the 4788 mutation. This PCR experiment uses a primer anchored in the predicted flanking sequence and the slp346 primer (anchored in the T-DNA insertion). Finding a PCR product of the size expected based on the sequence of the plasmid rescued clone confirms a valid rescue.
The sequence obtained from the above clone is used in an NCBI WWW blastn search against nucleotide sequence databases (Altschul et al. (1990) J Mol. Biol. 215: 403-410; Altschul et al (1997) Nucleic Acids Res. 25: 3389-3402). The search results show that the recovered sequence is identical to genomic DNA from *Arabidopsis* chromosome II BAC T9J22 (Genbank Acc. AC002505 (2739376)). The region of genomic DNA where the

insertion event occurred is annotated as encoding a putative permease (ACCESSION # 2739376). Primers are then designed to the 5' and 3' ends of predicted mRNA, and PCR is performed using DNA from an *Arabidopsis* cDNA library as the template. The resulting PCR product is TA-ligated and cloned (Original TA Cloning Kit, Invitrogen), and sequenced. The cDNA sequence is the same as the sequence predicted in the Genbank annotation, thus validating for the first time the putative open reading frame annotation.

Example 2: Expression of Recombinant 4788 Protein in *E. coli*

The coding region of the putative mature protein, corresponding to the cDNA clone, is subcloned into previously described expression vectors, and transformed into *E. coli* using the manufacturer's conditions. Specific examples include plasmids such as pBluescript (Stratagene, La Jolla, CA), pFLAG (International Biotechnologies, Inc., New Haven, CT), and pTrcHis (Invitrogen, La Jolla, CA).

Example 3:In vitro Recombination of 4788 Genes by DNA Shuffling

The *A. thaliana* 4788 gene encoding the 4788 protein is amplified by PCR. The resulting DNA fragment is digested by DNaseI treatment essentially as described (Stemmer et al. (1994) PNAS 91: 10747-10751) and the PCR primers are removed from the reaction mixture. A PCR reaction is carried out without primers and is followed by a PCR reaction with the primers, both as described (Stemmer et al. (1994) PNAS 91: 10747-10751). The resulting DNA fragments are cloned into pTRC99a (Pharmacia, Cat no: 27-5007-01) and transformed into *E.coli* uraA strain (Andersen et al. (1995) Journal of Bacteriol. 177(8): 2008-2013) by electroporation using the Biorad Gene Pulser and the manufacturer's conditions. The transformed bacteria are grown on medium that contains inhibitory concentrations of the inhibitor and those colonies that grow in the presence of the inhibitor are selected. Colonies that grow in the presence of normally inhibitory concentrations of inhibitor are picked and purified by repeated restreaking. Their plasmids are purified and the DNA sequences of cDNA inserts from plasmids that pass this test are then determined. In a similar reaction, PCR-amplified DNA fragments comprising the *A. thaliana* 4788 gene encoding the protein and PCR-amplified DNA fragments comprising the *E.coli* uraA gene are recombined *in vitro* and resulting variants with improved tolerance to the inhibitor are recovered as described above.

Example 4:In vitro Recombination of 4788 Genes by Staggered Extension Process

The *A. thaliana* 4788 gene encoding the 4788 protein and the *E.coli* uraA gene are each cloned into the polylinker of a pBluescript vector. A PCR reaction is carried out essentially as described (Zhao et al. (1998) Nature Biotechnology 16: 258-261) using the "reverse primer" and the "M13 -20 primer" (Stratagene Catalog). Amplified PCR fragments are digested with appropriate restriction enzymes and cloned into pTRC99a and mutated 4788 genes are screened as described in Example 3.

Example 5: InVitro Binding Assays

Recombinant 4788 protein is obtained, for example according to Example 2. The protein is immobilized on chips appropriate for ligand binding assays. The protein immobilized on the chip is exposed to sample compound in solution according to methods well know in the art. While the sample compound is in contact with the immobilized protein measurements capable of detecting protein-ligand interactions are conducted. Examples of such measurements are SELDI, biacore and FCS, described above. Compounds found to bind the protein are readily discovered in this fashion and are subjected to further characterization, for e.g. according to Example 6, below.

Example 6: Assay for Uracil Permease Activity

A simple assay is developed to screen for compounds that affect normal functioning of the 4788 activity. Such compounds are promising *in vitro* leads that can be tested for *in vivo* herbicidal activity. The assay consists of growing *E. coli uraA* harboring and functionally expressing the 4788 gene in minimal medium in the presence of a minimal inhibitory concentration of 5-fluorouracil. This is accomplished in a 96-well format for automated high-throughput screening. Compounds that are effective in blocking function of the 4788 protein inhibit the ability of the cells to take up 5-FU, and bacterial growth results. This growth is measured by simple turbidometric means. More specifically, the invention relates to a process of identifying an inhibitor of permease activity comprising:

a) introduce SEQ ID NO: 1 or SEQ ID NO:2, or nucleotide sequences substantially

similar thereto into a permease-deficient host cell, such as *E. coli uraA* such that said sequence is functionally expressable;

b) combining said host cell containing said nucleotide sequence, with a minimal inhibitory concentration of 5-fluorouracil, and with a compound to be tested for the ability to inhibit the activity of said permease, under conditions conducive to such inhibition,

c) measure host cell growth under the conditions of step (b); and

d) selecting said compound that inhibits host cell growth in step (c), and optionally

e) applying identified compound in step (d) to a plant to test for herbicidal activity, and

f) selecting compounds having herbicidal activity in step (e).

The above disclosed embodiments are illustrative. This disclosure of the invention will place one skilled in the art in possession of many variations of the invention. All such obvious and foreseeable variations are intended to be encompassed by the appended claims.

What Is Claimed Is:

1. A nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO:2.

2. The nucleotide sequence of claim 1, wherein the sequence encodes an amino acid sequence substantially similar to SEQ ID NO: 3.

3. The nucleotide sequence according to claim 1, wherein said nucleotide sequence is a plant nucleotide sequence.

4. The nucleotide sequence of claim 1, wherein the protein has permease activity.

5. The nucleotide sequence of claim 4, wherein the protein has purine or pyrimidine permease activity.

6. An amino acid sequence comprising an amino acid sequence encoded by a nucleotide sequence substantially similar to SEQ ID NO: 1 or SEQ ID NO: 2.

7. An amino acid sequence comprising an amino acid sequence substantially similar to SEQ ID NO:3.

8. The amino acid sequence of claim 6, wherein the protein has permease activity.

9. The amino acid sequence of claim 8, wherein the protein has purine or pyrimidine permease activity.

10. An amino acid sequence comprising at least 20 consecutive amino acid residues of the amino acid sequence encoded by SEQ ID NO: 1 or SEQ ID NO: 2.

11. An amino acid sequence comprising at least 20 consecutive amino acid residues of the amino acid sequence of SEQ ID NO: 3.

12. A chimeric gene comprising a promoter operatively linked to a nucleotide sequence according to anyone of claim 1 to 5.

13. A recombinant vector comprising a chimeric gene according to claim 12, wherein said vector is capable of being stably transformed into a host cell.

14. A host cell comprising a vector according to claim 13, wherein said nucleotide sequence is expressible in said cell.

15. A host cell according to claim 14, wherein said host cell is an eukaryotic cell.

16. A host cell according to claim 15, wherein said host cell is selected from the group consisting of an insect cell, a yeast cell, and a plant cell.

17. A host cell according to claim 15, wherein said host cell is a prokaryotic cell.

18. A plant or seed comprising a plant cell of claim 16.

19. A plant of claim 18, wherein said plant is tolerant to an inhibitor of permease activity.

20. A process for making nucleotides sequences encoding gene products having altered permease activity comprising,

a) shuffling a nucleotide sequence of claim 1,

b) expressing the resulting shuffled nucleotide sequences and

c) selecting for altered permease activity as compared to the permease activity of the gene product of said unmodified nucleotide sequence.

21. The process of claim 20, wherein the nucleotide sequence is SEQ ID NO: 1 or SEQ ID NO: 2.

22. The process of claim 20, wherein said permease activity is purine or pyrimidine permease activity.

23. A shuffled DNA molecule obtainable by the process of claim 20.

24. A shuffled DNA molecule obtained by the process of claim 20, wherein said shuffled DNA molecule encodes a gene product having enhanced tolerance to an inhibitor of permease activity.

25. A chimeric gene comprising a promoter operatively linked to a nucleotide sequence according to anyone of claims 23 to 24.

26. A recombinant vector comprising a chimeric gene according to claim 25, wherein said vector is capable of being stably transformed into a host cell.

27. A host cell comprising a vector according to claim 26, wherein said nucleotide sequence is expressible in said cell.

28. A host cell according to claim 27, wherein said host cell is an eukaryotic cell.

29. A host cell according to claim 27, wherein said host cell is selected from the group consisting of an insect cell, a yeast cell, and a plant cell.

30. A host cell according to claim 27, wherein said host cell is a prokaryotic cell.

31. A plant or seed comprising a plant cell of claim 29.

32. A plant of claim 31, wherein said plant is tolerant to an inhibitor of permease activity.

33. A process of identifying compounds having herbicidal activity comprising:

a) combining a protein of claim 6 and a compound to be tested for the ability to bind to said protein, under conditions conducive to binding,

b) selecting a compound identified in step (a) that is capable of binding said protein,

c) applying identified compound in step (b) to a plant to test for herbicidal activity,

and d) selecting compounds having herbicidal activity.

34. A compound having herbicidal activity identifiable according to the process of claim 33.

35. A process of identifying an inhibitor of permease activity having herbicidal activity comprising:

a) combining a permease and a compound to be tested for the ability to inhibit the activity of said permease, under conditions conducive to such inhibition,

b) selecting a compound identified in step (a) that is capable of inhibiting said permease activity,

c) applying identified compound in step (b) to a plant to test for herbicidal activity,

and d) selecting compounds having herbicidal activity.

36. The process of claim 35, wherein said permease is a purine or pyrimidine permease.

37. A compound having herbicidal activity identifiable according to the process of claim 35.

38. A process of identifying an inhibitor of permease activity comprising,

a) introducing SEQ ID NO: 1 or SEQ ID NO:2, or nucleotide sequences substantially similar thereto into a permease-deficient host cell, such as *E. coli uraA* such that said sequence is functionally expressable;

b) combining said host cell containing said nucleotide sequence, with a minimal inhibitory concentration of 5-fluorouracil, and with a compound to be tested for the ability to inhibit the activity of said permease, under conditions conducive to such inhibition,

c) measure host cell growth under the conditions of step (b); and

d) selecting said compound that inhibits host cell growth in step (c).

39. A compound having herbicidal activity identifiable according to the process of claim 38.

40. A method for suppressing the growth of a plant comprising, applying to said plant a compound that inhibits the activity of the amino acid sequence of claim 6 in an amount sufficient to suppress the growth of said plant.

41. The method of claim 40, wherein the compound is selected from the group consisting of the compounds of claim 37 and the compounds of claim 39.

42. A method of improving crops comprising, applying to a herbicide tolerant plant or seed selected from the group consisting of the plant or seed of claim 19 and the plant or seed of claim 32, a compound having herbicidal activity selected from the group consisting of the compounds of claim 34, the compounds of claim 37 and the compounds of claim 39, in an amount that inhibits the growth of undesired vegetation without significantly suppressing the growth of the herbicide tolerant plant or seed.

SEQUENCE LISTING

<110> Novartis AG

<120> Herbicide Target Gene and Methods

<130> PB/5-30638/CGC 2031

<140>
<141>
<150> Application No. 09/153,278
<151> 1998-09-15
<160> 4

<170> PatentIn Ver. 2.0

<210> 1
<211> 3710
<212> DNA
<213> Arabidopsis thaliana

<400> 1
tcaaacaagt gacttttat atttatctta ttaattctgt agtctcgaga ttcgcttttt 60

ctcctgctta cttcttttta tcatcttctc tttgtcggtt tgatttggaa aaactttctg 120

caaaaaaaag agagattctt aaacttttta tctcaagatc ctttcaaaaa atctgaaaag 180

agaagatttt taataaaaaa gaatggttga aactggtcac catcatcaac atccaccggc 240

accggctgca gccggtcatc cgccggttcc atccatggcg atggcgcgta acatgggaac 300

cacttggcct ccggcggagc aacttcatca tcttcaatat tgcatccact ctaacccttc 360

ttggcgttag tctctctctt tctgcttata ctactttctt tatttttgtt tcagatgtaa 420

aaaataaaaa ttataatctt ttcagtcaaa agttgtatta gttcagttct gtctctctct 480

ctctttatct gctttatgtt gtgtcttaat ttgtttacca aaagattagt ttagatctca 540

gtatctgatt cgattcttta tatttttcct ctgtgttgac tatattttag cttgtgttgg 600

ggactaatta gctttaaatc taatgataaa tgttgtataa ttgttcttga gacacactaa 660

ttgattaaga ttaaaataaa ctgtttcgtt attacttgtt tatgtagttg gtgaatccaa 720

aatgctataa gaattgaatc cttagcagtt aaaagtgtgg ggttgggttt atgctaacac 780

tggtcatgga aaatgaaatg caattagtgt tttaccatta caagaaatgc catttttaag 840

ggattttgta acaaaaggat tgtattattg tttcctttta agaattgatt tgtctcattt 900

ccattaacaa tgaatcattt tacagatgag acggttgtac tggctttcca gcattacatc 960

gttatgctcg ggactactgt tttgattgcc aacacactag tgtcaccaat gggtggagat 1020

cctgtatgtg atgtttttgg tttttggatc cggttgtcta atatttatgt tattatatga 1080

ttctgaaatc gaccgtttgg tattttgaac gataataggg tgataaagcg cgggttatcc 1140

agactatatt gtttatgtcg ggcataaaca cgttgctaca aacgcttatc gggacaaggc 1200

ttcccacggt tatgggagta tcgttcgcgt atgttcttcc cgtattgtct ataatcagag 1260

attacaacaa tggtcaattc gattccgaga aacaggttaa agctcgtttt atacttgcgc 1320

ggtacggttg ctatatatat attttttttt ttacttgaca aatgccattt tccttgtaga 1380

gattccgaca tactatgaga acggtacaag gatcattaat catttcttca ttcgtcaaca 1440

tcataatcgg atatggtcag gcatggggga acttaataag gtaatataaa agatcaatat 1500

ataacattac tagtttagaa gctgatggaa gttttaacat ttaatgatta tggtttcttt 1560

aatgcagaat ctttagtccc atcattgttg tgcccgttgt ttctgttgtg agccttggcc 1620

tattccttag aggcttccca ctggtaacaa tttcaactaa agtattcagc tttaacaaat 1680

tttaactcct tcactaagtc attatgtgat ctcttcggat tttaacagct tgcaaactgt 1740

gtggaaatcg gtctaccaat gctgattctg ttgatcatca cacagcaagt cgggcttttt 1800

attattttcc gcgctatata tggtggtgat tgtttagata tgtaatatta aatcttgcta 1860

ctgttcatct tgcagtatct taaacatgca ttctcaagga tttctatgat tcttgaaaga 1920

tatgctttac ttgtttgcct ggctatcata tgggcttttg ctgctatcct tactgtttct 1980

ggtgcttata ataatgtttc taccgcaaca aaacaaagtt gtcgaactga tcgtgccttt 2040

cttatgtcat cagctccctg gtaggttgat tactatttga cttgattctt cttttcttct 2100

aaggtctaaa tatctttata ctcagtcact caaactcatc tttcaggatt agaatcccat 2160

atccattcca gtgggggact ccgatattca aagcgagtca tgtttttgga atgtttggtg 2220

ctgcaattgt cgcatctgca gaggttgtta ttcatcatat tagtttctag cttatgttaa 2280

atattcttca tgttacatgc tacaacttgt tcttatgtta tgtatatcct tatttttctt 2340

agtctaccgg agtatttttc gctgcatcta gactagcagg agcaacagcg cctccagcac 2400

atgtcgtctc tcgtagtatc ggtctacagg ttttatttcc agacactaag aaagtttttt 2460

ttaatctttt cgttttctgt tctctctgct aatgttccga ataagtctat aagctgttat 2520

attttccttt aagggtattg gtgtgctcct tgaaggaata tttggttcca ttactggcaa 2580

taccgcgtcc gtgtaagttc taaatatctc ttgctatatg tgctactatc ctttcagaat 2640

ttatacaaag aaactaggta tataattcat cttgatgata tatacaggga aaatgtcggt 2700

ctccttggcc tcacacgaat agggagtaga cgagtggtgc aggtttcaac gtttttcatg 2760

atatttttct ccatatttgg tttgtctttc aaccctctaa tcagtcatct tgactaagta 2820

tagaaagtag ccgttcatgg gttttaaatc cgcgtgtttt caatgatctt caggaaaatt 2880

tggcgcgttc tttgcgtcta ttccgcttcc aatctttgca ggcgtatact gtatactact 2940

tggaatcgtt ggtgagtagc attttgtata tgaacctaca ctacctgatt tcttgtaaaa 3000

gccggggctg cttatgatac ttttgttatg tgattggccc agttgctgtt ggaatatcgt 3060

ttatacagtt taccgacact aattcgatga gaaacatgta cgtcattggt gtctctctct 3120

tcttaagtct ttccatcgct cagtactttc ttgccaacac ttcaagagca ggatatggac 3180

cagttaggac agcaggagga tgggtaagcc tttcaaagaa ccattgtttg aaacaccatt 3240

ttacggtagt atagggagtg atataatatt tctactatat agtgtttctt tttcttaaat 3300

gtgatgtcgc ggtgaattgt ggtgcagttc aacgatatac ttaatacgat atttgcttcg 3360

gctccgttgg tggcgaccat tcttgcgacc atactagata acacgttgga agcaagacat 3420

gcaagtgacg acgcaagagg aatcccgtgg tggaagccct tccagcacag gaacggagac 3480

ggcaggaacg atgagttcta tagtatgccc cttagaatca acgagttaat gccgacacgg 3540

ttcctttgaa gactgcccct gaacgtttct tctgtatttg gaaatgtaag atatgattat 3600

gtgcatacct tgtagcttca ttggggaaaa attgagtcca gtggatacaa atgaacatag 3660

gcctttgatg gaaaaagcta ttttttgca aactatataa cttgtgttac 3710

<210> 2
<211> 1674
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (4)..(1659)

<400> 2

aga atg gtt gaa act ggt cac cat cat caa cat cca ccg gca ccg gct 48
Met Val Glu Thr Gly His His His Gln His Pro Pro Ala Pro Ala
1 5 10 15

gca gcc ggt cat ccg ccg gtt cca tcc atg gcg atg gcg cgt aac atg 96
Ala Ala Gly His Pro Pro Val Pro Ser Met Ala Met Ala Arg Asn Met
20 25 30

gga acc act tgg cct ccg gcg gag caa ctt cat cat ctt caa tat tgc 144
Gly Thr Thr Trp Pro Pro Ala Glu Gln Leu His His Leu Gln Tyr Cys
35 40 45

atc cac tct aac cct tct tgg cat gag acg gtt gta ctg gct ttc cag 192
Ile His Ser Asn Pro Ser Trp His Glu Thr Val Val Leu Ala Phe Gln
50 55 60

cat tac atc gtt atg ctc ggg act act gtt ttg att gcc aac aca cta 240
His Tyr Ile Val Met Leu Gly Thr Thr Val Leu Ile Ala Asn Thr Leu
65 70 75

gtg tca cca atg ggt gga gat cct ggt gat aaa gcg cgg gtt atc cag 288
Val Ser Pro Met Gly Gly Asp Pro Gly Asp Lys Ala Arg Val Ile Gln
80 85 90 95

```
act ata ttg ttt atg tcg ggc ata aac acg ttg cta caa acg ctt atc   336
Thr Ile Leu Phe Met Ser Gly Ile Asn Thr Leu Leu Gln Thr Leu Ile
                100                 105                 110

ggg aca agg ctt ccc acg gtt atg gga gta tcg ttc gcg tat gtt ctt   384
Gly Thr Arg Leu Pro Thr Val Met Gly Val Ser Phe Ala Tyr Val Leu
            115                 120                 125

cct gta ttg tct ata atc aga gat tac aac aat ggt caa tcc gat tcc   432
Pro Val Leu Ser Ile Ile Arg Asp Tyr Asn Asn Gly Gln Ser Asp Ser
        130                 135                 140

gag aaa cag aga ttc cga cat act atg aga acg gta caa gga tca tta   480
Glu Lys Gln Arg Phe Arg His Thr Met Arg Thr Val Gln Gly Ser Leu
    145                 150                 155

atc att tct tca ttc gtc aac atc ata atc gga tat ggt cag gca tgg   528
Ile Ile Ser Ser Phe Val Asn Ile Ile Ile Gly Tyr Gly Gln Ala Trp
160                 165                 170                 175

ggg aac tta ata aga atc ttt agt ccc atc att gtt gtg ccc gtt gtt   576
Gly Asn Leu Ile Arg Ile Phe Ser Pro Ile Ile Val Val Pro Val Val
                180                 185                 190

tct gtt gtg agc ctt ggc cta ttc ctt aga ggc ttc cca ctg ctt gca   624
Ser Val Val Ser Leu Gly Leu Phe Leu Arg Gly Phe Pro Leu Leu Ala
            195                 200                 205

aac tgt gtg gaa atc ggt cta cca atg ctg att ctg ttg atc atc aca   672
Asn Cys Val Glu Ile Gly Leu Pro Met Leu Ile Leu Leu Ile Ile Thr
        210                 215                 220

cag caa tat ctt aaa cat gca ttc tca agg att tct atg att ctt gaa   720
Gln Gln Tyr Leu Lys His Ala Phe Ser Arg Ile Ser Met Ile Leu Glu
    225                 230                 235

aga tat gct tta ctt gtt tgc ccg gct atc ata tgg gct ttt gct gct   768
Arg Tyr Ala Leu Leu Val Cys Pro Ala Ile Ile Trp Ala Phe Ala Ala
240                 245                 250                 255

atc ctt act gtt tct ggt gct tat aat aat gtt tct acc gca aca aaa   816
Ile Leu Thr Val Ser Gly Ala Tyr Asn Asn Val Ser Thr Ala Thr Lys
                260                 265                 270

caa agt tgt cga acg gat cgt gcc ttt ctt atg tca tca gct ccc tgg   864
Gln Ser Cys Arg Thr Asp Arg Ala Phe Leu Met Ser Ser Ala Pro Trp
            275                 280                 285

att aga atc cca tat cca ttc cag tgg ggg act ccg ata ttc aaa gcg   912
Ile Arg Ile Pro Tyr Pro Phe Gln Trp Gly Thr Pro Ile Phe Lys Ala
        290                 295                 300

agt cat gtt ttt gga atg ttt ggt gct gca att gtc gca tct gca gag   960
Ser His Val Phe Gly Met Phe Gly Ala Ala Ile Val Ala Ser Ala Glu
    305                 310                 315

tct acc gga gta ttt ttc gct gca tct aga tta gca gga gca aca gcg  1008
Ser Thr Gly Val Phe Phe Ala Ala Ser Arg Leu Ala Gly Ala Thr Ala
320                 325                 330                 335

cct cca gca cat gtc gtc tct cgt agt atc ggt cta cag ggt att ggt  1056
```

```
Pro Pro Ala His Val Val Ser Arg Ser Ile Gly Leu Gln Gly Ile Gly
                340                 345                 350

gtg ctc ctt gaa gga ata ttt ggt tcc att act ggc aat acc gcg tcc   1104
Val Leu Leu Glu Gly Ile Phe Gly Ser Ile Thr Gly Asn Thr Ala Ser
            355                 360                 365

gtg gaa aat gtc ggt ctc ctt ggc ctc gca cga ata ggg agt aga cga   1152
Val Glu Asn Val Gly Leu Leu Gly Leu Ala Arg Ile Gly Ser Arg Arg
        370                 375                 380

gtg gtg cag gtt tca acg ttt ttc atg ata ttt ttc tcc ata ttt gga   1200
Val Val Gln Val Ser Thr Phe Phe Met Ile Phe Phe Ser Ile Phe Gly
    385                 390                 395

aaa ttt ggc gcg ttc ttt gcg tct att ccg ctt cca atc ttt gca ggc   1248
Lys Phe Gly Ala Phe Phe Ala Ser Ile Pro Leu Pro Ile Phe Ala Gly
400                 405                 410                 415

ata tac tgt ata cta ctt gga atc gtt gtt gct gtt gga ata tcg ttt   1296
Ile Tyr Cys Ile Leu Leu Gly Ile Val Val Ala Val Gly Ile Ser Phe
                420                 425                 430

ata cag ttt acc gac act aat tcg atg aga aac atg tac gtc att ggt   1344
Ile Gln Phe Thr Asp Thr Asn Ser Met Arg Asn Met Tyr Val Ile Gly
            435                 440                 445

gtc tct ctc ttc tta agt ctt tcc atc gct cag tac ttt ctt gcc aac   1392
Val Ser Leu Phe Leu Ser Leu Ser Ile Ala Gln Tyr Phe Leu Ala Asn
        450                 455                 460

act tca aga gca gga tat gga cca gtt agg aca gca gga gga tgg ttc   1440
Thr Ser Arg Ala Gly Tyr Gly Pro Val Arg Thr Ala Gly Gly Trp Phe
    465                 470                 475

aac gat ata ctt aat acg ata ttt gct tcg gct ccg ttg gtg gcg acc   1488
Asn Asp Ile Leu Asn Thr Ile Phe Ala Ser Ala Pro Leu Val Ala Thr
480                 485                 490                 495

att ctt gcg acc ata cta gat aac acg ttg gaa gca aga cat gca agt   1536
Ile Leu Ala Thr Ile Leu Asp Asn Thr Leu Glu Ala Arg His Ala Ser
                500                 505                 510

gac gac gca aga gga atc ccg tgg tgg aag ccc ttc cag cac agg aac   1584
Asp Asp Ala Arg Gly Ile Pro Trp Trp Lys Pro Phe Gln His Arg Asn
            515                 520                 525

gga gac ggc agg aac gat gag ttc tat agt atg ccc ctt aga atc aac   1632
Gly Asp Gly Arg Asn Asp Glu Phe Tyr Ser Met Pro Leu Arg Ile Asn
        530                 535                 540

gag tta atg ccg aca cgg ttc ctt tga agactgcccc tgaac              1674
Glu Leu Met Pro Thr Arg Phe Leu
    545                 550
```

<210> 3
<211> 551
<212> PRT
<213> Arabidopsis thaliana

<400> 3

```
Met Val Glu Thr Gly His His His Gln His Pro Pro Ala Pro Ala Ala
 1               5                  10                  15

Ala Gly His Pro Pro Val Pro Ser Met Ala Met Ala Arg Asn Met Gly
            20                  25                  30

Thr Thr Trp Pro Pro Ala Glu Gln Leu His His Leu Gln Tyr Cys Ile
        35                  40                  45

His Ser Asn Pro Ser Trp His Glu Thr Val Val Leu Ala Phe Gln His
    50                  55                  60

Tyr Ile Val Met Leu Gly Thr Thr Val Leu Ile Ala Asn Thr Leu Val
65                  70                  75                  80

Ser Pro Met Gly Gly Asp Pro Gly Asp Lys Ala Arg Val Ile Gln Thr
                85                  90                  95

Ile Leu Phe Met Ser Gly Ile Asn Thr Leu Leu Gln Thr Leu Ile Gly
            100                 105                 110

Thr Arg Leu Pro Thr Val Met Gly Val Ser Phe Ala Tyr Val Leu Pro
        115                 120                 125

Val Leu Ser Ile Ile Arg Asp Tyr Asn Asn Gly Gln Ser Asp Ser Glu
    130                 135                 140

Lys Gln Arg Phe Arg His Thr Met Arg Thr Val Gln Gly Ser Leu Ile
145                 150                 155                 160

Ile Ser Ser Phe Val Asn Ile Ile Ile Gly Tyr Gly Gln Ala Trp Gly
                165                 170                 175

Asn Leu Ile Arg Ile Phe Ser Pro Ile Ile Val Val Pro Val Val Ser
            180                 185                 190

Val Val Ser Leu Gly Leu Phe Leu Arg Gly Phe Pro Leu Leu Ala Asn
        195                 200                 205

Cys Val Glu Ile Gly Leu Pro Met Leu Ile Leu Leu Ile Ile Thr Gln
    210                 215                 220

Gln Tyr Leu Lys His Ala Phe Ser Arg Ile Ser Met Ile Leu Glu Arg
225                 230                 235                 240

Tyr Ala Leu Leu Val Cys Pro Ala Ile Ile Trp Ala Phe Ala Ala Ile
                245                 250                 255

Leu Thr Val Ser Gly Ala Tyr Asn Asn Val Ser Thr Ala Thr Lys Gln
            260                 265                 270

Ser Cys Arg Thr Asp Arg Ala Phe Leu Met Ser Ser Ala Pro Trp Ile
        275                 280                 285

Arg Ile Pro Tyr Pro Phe Gln Trp Gly Thr Pro Ile Phe Lys Ala Ser
    290                 295                 300

His Val Phe Gly Met Phe Gly Ala Ala Ile Val Ala Ser Ala Glu Ser
305                 310                 315                 320
```

Thr Gly Val Phe Phe Ala Ala Ser Arg Leu Ala Gly Ala Thr Ala Pro
325 330 335

Pro Ala His Val Val Ser Arg Ser Ile Gly Leu Gln Gly Ile Gly Val
340 345 350

Leu Leu Glu Gly Ile Phe Gly Ser Ile Thr Gly Asn Thr Ala Ser Val
355 360 365

Glu Asn Val Gly Leu Leu Gly Leu Ala Arg Ile Gly Ser Arg Arg Val
370 375 380

Val Gln Val Ser Thr Phe Phe Met Ile Phe Phe Ser Ile Phe Gly Lys
385 390 395 400

Phe Gly Ala Phe Phe Ala Ser Ile Pro Leu Pro Ile Phe Ala Gly Ile
405 410 415

Tyr Cys Ile Leu Leu Gly Ile Val Val Ala Val Gly Ile Ser Phe Ile
420 425 430

Gln Phe Thr Asp Thr Asn Ser Met Arg Asn Met Tyr Val Ile Gly Val
435 440 445

Ser Leu Phe Leu Ser Leu Ser Ile Ala Gln Tyr Phe Leu Ala Asn Thr
450 455 460

Ser Arg Ala Gly Tyr Gly Pro Val Arg Thr Ala Gly Gly Trp Phe Asn
465 470 475 480

Asp Ile Leu Asn Thr Ile Phe Ala Ser Ala Pro Leu Val Ala Thr Ile
485 490 495

Leu Ala Thr Ile Leu Asp Asn Thr Leu Glu Ala Arg His Ala Ser Asp
500 505 510

Asp Ala Arg Gly Ile Pro Trp Trp Lys Pro Phe Gln His Arg Asn Gly
515 520 525

Asp Gly Arg Asn Asp Glu Phe Tyr Ser Met Pro Leu Arg Ile Asn Glu
530 535 540

Leu Met Pro Thr Arg Phe Leu
545 550

<210> 4
<211> 20
<212> DNA
<213> Arabidopsis thaliana

<400> 4
gcggacatct acatttttga 20

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/82, 9/00, A01H 5/00, C12Q 1/68 | A3

(11) International Publication Number: **WO 00/15809**

(43) International Publication Date: 23 March 2000 (23.03.00)

(21) International Application Number: PCT/EP99/06766

(22) International Filing Date: 13 September 1999 (13.09.99)

(30) Priority Data:
09/153,278 15 September 1998 (15.09.98) US

(71) Applicant *(for all designated States except AT US)*: NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH–4058 Basel (CH).

(71) Applicant *(for AT only)*: NOVARTIS–ERFINDUNGEN VERWALTUNGSGESELLSCHAFT M.B.H. [AT/AT]; Brunner Strasse 59, A–1230 Vienna (AT).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LEVIN, Joshua, Zvi [US/US]; 1008 Urban Avenue, Durham, NC 27701 (US). POTTER, Sharon, Lee [US/US]; 3837 Whispering Branch Road, Raleigh, NC 27613 (US). BUDZISZEWSKI, Gregory, Joseph [US/US]; 2016 Englewood Avenue, Durham, NC 27705 (US). WARD, Eric, Russell [US/US]; 3761 Bentley Drive, Durham, NC 27707 (US). BERNASCONI, Paul [US/US]; 102 Windhover Drive, Chapel Hill, NC 27514 (US).

(74) Agent: BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent & Trademark Dept., CH–4002 Basel (CH).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
29 June 2000 (29.06.00)

(54) Title: URACIL PERMEASE FROM ARABIDOPSIS AS HERBICIDAL TARGET GENE

(57) Abstract

The invention relates to a gene isolated from *Arabidopsis* that codes for a protein essential for seedling growth. The invention also includes the methods of using this protein to discover new herbicides, based on the essentiality of the gene for normal growth and development. The invention can also be used in a screening assay to identify inhibitors that are potential herbicides. The invention is also applied to the development of herbicide tolerant plants, plant tissues, plant seeds, and plant cells.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/EP 99/06766

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/82 C12N9/00 A01H5/00 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ROUNSLEY, S.D., ET AL.: "Arabidopsis thaliana chromosome II BAC T9J22 sequence, complete sequence" EMBL ACCESSION NO:AC002505, 11 September 1997 (1997-09-11), XP002132053 cited in the application	1-17
Y	see nts 59581-63290 -& "PUTATIVE PERMEASE." TREMBL ACCESSION NO:O48718, 1 June 1998 (1998-06-01), XP002132054 & "Xanthine/uracil permeases family" PFAM DATABASE ACCESSION NO:PF00860,	12-17, 33,35
Y	DE 43 43 527 A (SCHERING AG) 22 June 1995 (1995-06-22) the whole document	12-17, 33,35

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 March 2000	17/03/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	MADDOX, A

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	DBEST ID:1436641, 5 January 1998 (1998-01-05), XP002132055 & "21525 CD4-15 Arabidopsis thaliana cDNA clone G3C2T7." EMBL ACCESSION NO:N96722, 19 April 1996 (1996-04-19),	1-17
X	DBEST ID:32834, 27 September 1993 (1993-09-27), XP002132056 & "A. thaliana transcribed sequence; clone YAY053; 5' end." EMBL ACCESSION NO:Z26369, 23 September 1993 (1993-09-23),	1-17
X	DBEST ID:43468,15 March 1994 (1994-03-15), XP002132057 & " A. thaliana transcribed sequence; clone VBVAA01." EMBL ACCESSION NO:Z30726, 14 March 1994 (1994-03-14),	1-17
X	SCHULTES, N.P, ET AL.: "Leaf permeasel gene of maize is required for chloroplast development" THE PLANT CELL, vol. 8, 1996, pages 463-475, XP002132058 cited in the application the whole document & EMBL ACCESSION NO:U43034, 13 January 1996 (1996-01-13), & TREMBL ACCESSION NO:Q41760, 1 November 1996 (1996-11-01),	1-10
A	WO 95 25114 A (UNIV TENNESSEE RES CORP) 21 September 1995 (1995-09-21) claim 16	19
A	WO 98 31837 A (DELCARDAYRE STEPHEN B ;MAXYGEN INC (US); MINSHULL JEREMY (US); NES) 23 July 1998 (1998-07-23) page 28, line 19 -page 31, line 37	20-22
A	WO 97 35966 A (MAXYGEN INC ;MINSHULL JEREMY (US); STEMMER WILLEM P C (US)) 2 October 1997 (1997-10-02) the whole document	20-22

-/--

4

Int. Jonal Application No
PCT/EP 99/06766

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passagee	Relevant to claim No.
A	ABELL ET AL: "Biochemical aproaches to herbicide discovery: advances in enzyme target identification and inhibitor design" WEED SCIENCE,US,WEED SCIENCE SOCIETY OF AMERICA, CHAMPAIGN, IL, vol. 44, 1996, pages 734-742, XP002094170 ISSN: 0043-1745 the whole document	33,35, 36,38
A	DATABASE BIOSIS 'Online! BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US1978 LOSSON R ET AL: "PROPERTIES OF 3 DISTINCT PYRIMIDINE TRANSPORT SYSTEMS IN YEAST EVIDENCE FOR DISTINCT ENERGY COUPLING" Database accession no. PREV197967031009 XP002132060 abstract & BIOCHIMICA ET BIOPHYSICA ACTA 1978 * EN *, vol. 513, no. 2, 1978, pages 296-300, ISSN: 0006-3002	33,35, 36,38

INTERNATIONAL SEARCH REPORT

...ernational application No.
PCT/EP 99/06766

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 23-32,34,37,39-42
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/EP 99 /06766

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 23-32,34,37,39-42

Claims 23-32,34,37,39-42 refer to DNA molecules and compounds without giving a true technical characterization. Moreover no such compounds are defined in the application. In consequence, the scope of said claims is amibiguous and vague, and thier subject-matter is not sufficiently disclosed and suppported (Art. 5 and 6) PCT. No search can be carried out for such purely speculative claims whose wording is in fact a mere recitation of the results to be achieved

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. ..onal Application No
PCT/EP 99/06766

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
DE 4343527 A	22-06-1995	AU 701626 B	04-02-1999
		AU 1314895 A	03-07-1995
		CA 2179062 A	22-06-1995
		WO 9516913 A	22-06-1995
		EP 0734525 A	02-10-1996
		HU 75077 A	28-04-1997
		JP 9506512 T	30-06-1997
		US 5750362 A	12-05-1998
WO 9525114 A	21-09-1995	US 5689039 A	18-11-1997
		CA 2207851 A	21-09-1995
		EP 0767797 A	16-04-1997
		JP 10500562 T	20-01-1998
WO 9831837 A	23-07-1998	AU 5920998 A	07-08-1998
WO 9735966 A	02-10-1997	US 5837458 A	17-11-1998
		AU 2337797 A	17-10-1997
		AU 2542697 A	17-10-1997
		CA 2247930 A	02-10-1997
		EP 0932670 A	04-08-1999
		EP 0906418 A	07-04-1999
		WO 9735957 A	02-10-1998
		AU 713952 B	16-12-1999
		AU 1087397 A	19-06-1997
		CA 2239099 A	05-06-1997
		EP 0876509 A	11-11-1998
		EP 0911396 A	28-04-1999
		WO 9720078 A	05-06-1997